ERGODIC THEORY AND THE DUALITY PRINCIPLE ON HOMOGENEOUS SPACES

ALEXANDER GORODNIK AND AMOS NEVO

Abstract. We prove mean and pointwise ergodic theorems for the action of a discrete lattice subgroup in a connected algebraic Lie group on infinite volume homogeneous algebraic varieties. Under suitable necessary conditions, our results are quantitative, namely we establish rates of convergence in the mean and pointwise ergodic theorems, which can be estimated explicitly. Our results give a precise and in most cases optimal quantitative form to the duality principle governing dynamics on homogeneous spaces. We illustrate their scope in a variety of equidistribution problems.

Contents

1. Introduction

1.1. **Ergodic theory and the duality principle on homogeneous spaces.** The classical framework of ergodic theory usually includes a compact space X equipped with finite measure and an action of a countable group Γ which preserves this measure. In order to study the distribution of the orbits $x\Gamma$ in X, one chooses an increasing sequence $\{\Gamma_t\}_{t\ge t_0}$ of finite subsets of Γ and considers the averaging operators

$$\pi_X(\lambda_t)\phi(x) = \frac{1}{|\Gamma_t|}\sum_{\gamma\in\Gamma_t}\phi(x\gamma),$$

Date: March 4, 2013.
The first author was supported in part by EPSRC, ERC, and RCUK.
The second author was supported by ISF grant.

defined for functions ϕ on X. One of the fundamental problems in ergodic theory is to understand the asymptotic behaviour of $\pi_X(\lambda_t)\phi$ as $t \to \infty$. This question has been studied extensively when Γ is an amenable group and the averages are supported on Følner sets (see [Ne] for a survey, and [AAB] for a detailed recent discussion). Subsequently pointwise ergodic theorems were established for some classes of nonamenable groups, including lattice subgroups in semisimple algebraic groups, with the averages supported on norm balls (see [GN1] for a comprehensive discussion).

The situation when X is a non-compact locally compact space equipped with an infinite Radon measure is also of great interest, but it involves new highly nontrivial challenges. Indeed, in this case, the averages $\pi_X(\lambda_t)\phi$ considered above typically converge to zero. In order to obtain significant information about the distribution of orbits, it is natural to introduce the (normalized) orbit-sampling operators

$$\pi_X(\lambda_t)\phi(x) = \frac{1}{V(t)} \sum_{\gamma\in\Gamma_t} \phi(x\gamma),$$

where $V(t)$ is a suitable normalization, which one would like to choose so as to guarantee that the limit as $t \to \infty$ exists and is nontrivial. While it is well-known [Aa, Th. 2.4.2])) that for an action of a single transformation no such normalization exists, we shall achieve this objective for an extensive family of actions of lattices on infinite-volume homogeneous spaces, and proceed to develop a systematic and quantitative ergodic theory for the operators $\pi_X(\lambda_t)$. Our general results describe, in particular, the distribution of lattice orbits on the de-Sitter space, answering questions raised by Arnol'd [Ar, 1996-15, 2002-16].

The methods that we develop in order to obtain this goal are of very general nature and amount to establishing a quantitative form of an abstract duality principle for homogeneous spaces. Namely, if $X \simeq H\backslash G$ is a homogeneous space of a locally compact second countable group G and Γ is a discrete lattice subgroup in G, we reduce the ergodic-theoretic properties of the Γ-orbits on X to the ergodic-theoretic properties of the H-orbits in the (dual) action of H on G/Γ.

We will develop below an axiomatic framework in which the quantitative duality principle will be established in full generality (Sections 2-7). Our principal motivation for taking an abstract approach is the fact that the present paper does not exhaust the range of validity and the diverse applications of the quantitative ergodic theorems that we develop. Most importantly, essentially all of our arguments carry over with minor modifications to the case of general S-algebraic groups over fields of characteristic zero. We also note that many of our arguments carry over to homogeneous spaces of adele groups, as well as to S-algebraic groups over fields of positive characteristic. To illustrate their utility, we refer to [GGN] for an application of quantitative ergodic duality arguments to Diophantine approximation on homogeneous algebraic varieties in the S-algebraic set-up, answering some long-standing questions raised originally by S. Lang [La]. We plan to return to the quantitative duality principle and its applications in the context of homogenous spaces of S-algebraic groups in the future, but in the interest of brevity will confine ourselves in the present paper to connected Lie groups and their homogeneous spaces.

Let us now note that the subject of ergodic theory of non-amenable groups acting on infinite-measure spaces is full of surprises and exhibits several remarkable features which do not arise in the classical case of amenable groups acting on probability spaces. Let us mention the following ones.

(1) As already noted, the very existence of a normalization $V(t)$ for the orbit-sampling operators $\pi_X(\lambda_t)$ is impossible in the case of $\mathbb{Z}$-actions ; but we will also encounter the remarkable phenomenon that the growth of the sampling sets Γ_t may be exponential in t, while the normalization $V(t)$ is polynomial in t. Thus, for a point x in a given bounded set $D \subset X = H \setminus G$, the set of return points $x \cdot \Gamma_t \cap D$ is *logarithmic* in the size of $x \cdot \Gamma_t$, and yet the set of return points is almost surely equidistributed in D. In general, the set of return points will be exponentially small compared to the set of orbit points.

(2) The ergodic theorems we prove assert that under suitable conditions the averages $\pi_X(\lambda_t)$ converge in a suitable sense to a limiting distribution :

$$\lim_{t\to\infty} \pi_X(\lambda_t)\phi(x) = \int_X \phi \, d\nu_x \, .$$

However, the limiting distribution may fail to be invariant under the Γ-action, and may depend non-trivially on the initial point x, exhibiting distinctly non-amenable phenomena.

(3) When the dual action of H on G/Γ has a suitable spectral gap, we will establish an effective rate of convergence of the orbit-sampling operators to the limiting distribution, in the mean and sometimes pointwise. As a consequence, we will obtain quantitative ergodic theorems for actions on homogeneous spaces. Again these are new and distinctly non-amenable phenomena.

We remark that the ergodic theorems we establish have another significant set of applications which involves ratio ergodic theorems on homogeneous spaces, a subject raised originally by Kazhdan [K] for the Euclidean group. We will state some ratio ergodic theorems and comment further on this subject below.

Before turning to the exact statements of our main results in the next section, let us make one further comment on their scope. As explained in [GN1], the ergodic theory of non-amenable groups has to contend with the absence of asymptotic invariance and transference arguments that play a pivotal role in amenable ergodic theory. An indispensable tool to compensate for this absence is the existence of detailed quantitative volume estimates for the sampling sets involved in our analysis. These estimates include quantitative volume asymptotics, as well as quantitative stability and regularity properties, which will be explained in detail and exploited in our analysis below. The verification of these volume estimates is an intricate and challenging task which played a central role in [GN1]. Here we elaborate on it further to the extent required to establish our principal objective, which is the systematic development of ergodic theory for lattice subgroups of algebraic groups acting on homogeneous algebraic varieties. In principle, our results hold whenever the required volume estimates are valid, and there are grounds to expect that such volume estimates may be satisfied beyond the case of algebraic groups. However,

the volume estimates are definitely not valid for completely general Lie groups and their homogeneous manifolds, as demonstrated in [GW, 12.2]. For this reason, we will restrict the discussion to algebraic groups acting on algebraic homogeneous spaces, and with the sampling sets being defined by a homogeneous polynomial, or in some cases, a norm.

1.2. **Statement of the main results.** Let us start be introducing notation that will be in force throughout the paper. Let $G \subset \mathrm{SL}_d(\mathbb{R})$ be a connected closed subgroup. Let $H \subset G$ be a closed subgroup, let $X = H \setminus G$ be the corresponding homogeneous space, and let Γ be a discrete lattice in G.

For a proper function P positive except at 0, we consider the family of finite sets $\Gamma_t = \{\gamma \in \Gamma : \log P(\gamma) \le t\}$. Our main object of study will be the associated orbit-sampling operators $\sum_{\gamma\in\Gamma_t} \phi(x\gamma)$ with $x \in X$ and $\phi : X \to \mathbb{R}$, whose properties reflect the distribution of the orbits of Γ in X. We will use normalization functions $V(t)$ of two kinds for the orbit sampling operators. The first is defined intrinsically:

$$\pi_X(\lambda_t)\phi(x) = \frac{1}{\mathrm{vol}(H_t)} \sum_{\gamma\in\Gamma_t} \phi(x\gamma),$$

where $H_t = \{h \in H : \log P(h) \le t\}$. and the second reflects our knowledge of the volume asymptotics of H_t (when applicable) :

$$\pi_X(\tilde{\lambda}_t)\phi(x) = \frac{1}{e^{at}t^b} \sum_{\gamma\in\Gamma_t} \phi(x\gamma)$$

with $a \ge 0$ and $b \ge 0$. In the context of algebraic groups with the sets H_t defined by the homogeneous polynomial P, the volume of H_t does indeed have $c\, e^{at}t^b$ as its main term, so that the two operators are comparable.

We fix a smooth measure ξ on X with strictly positive density. Our discussion below will focus on a fixed (but arbitrary) compact set $D \subset X$. We assume that $D = \overline{\mathrm{Int}\,(D)}$ and that the boundary of $\mathrm{Int}\,(D)$ has zero measure, and call D a compact domain in this case. we denote by $L^p(D)$ the space of L^p-integrable functions ϕ with $\mathrm{supp}(\phi) \subset D$ equipped with the norm

$$\|\phi\|_{L^p(D)} = \left(\int_D |\phi|^p \, d\xi\right)^{1/p}.$$

Since the space $L^p(D)$ does not depend on the measure ξ and different ξ's lead to equivalent norms, we suppress ξ from the notation. We also denote by $L^p_l(D)$ the space of Sobolev function with support in D (see Section 2) and by $L^p_l(D)^+$ the subset of $L^p_l(D)$ consisting of nonnegative functions.

Our main results are formulated in the three theorems stated below. We consider three possibilities for the structure of the stability group H and the volume growth of the sets H_t, as measured by our choice of Haar measure on H. As noted above, we will require stringent regularity conditions on H_t and its volume, and for this reason will assume from now on that G and H are almost algebraic groups of $\mathrm{SL}_d(\mathbb{R})$ (namely they are of finite index in their Zariski closure over $\mathbb{R}$), and that P is a homogeneous polynomial on the linear space $\mathrm{Mat}_d(\mathbb{R})$.

The first main result, Theorem 1.1, deals with case of where the volume growth of H_t in polynomial in t, so that the normalization factor $V(t)$ of the sampling operators is polynomial as well. As we shall see below in Lemma 8.2, this forces H to be isomorphic to the almost direct product of an $\mathbb{R}$-diagonalizable torus and a compact group. Theorem 1.1 establish a mean and pointwise ergodic theorem for the normalized sampling operators $\pi_X(\tilde{\lambda}_t)$ in every L^p-space, $1 \le p < \infty$. Under the further assumptions that G is semisimple, the lattice Γ is irreducible, and the action of G on $L_0^2(G/\Gamma)$ has strong spectral gap, it establishes a pointwise ergodic theorem with a polynomial rate of convergence for Sobolev functions. We recall that a unitary representation of a connected semisimple group is said to have a *strong spectral gap* if its restriction to every simple factor L is isolated from the trivial representation of L.

Our second main result, Theorem 1.2, deals with the case where G is semisimple, the volume growth of H_t is exponential, and allows H to be either amenable or non-amenable. It establishes a mean ergodic theorems for the normalized sampling operators in L^p, $1 \le p < \infty$, as well as a pointwise ergodic theorem for Sobolev functions, and a pointwise ergodic theorem with a rate of convergence when the functions are subanalytic.

Our third main result, Theorem 1.3, assumes that of the underlying algebraic groups G and H, at least one is semisimple, that the lattice Γ is irreducible and that the stability group H is non-amenable subgroup which is non-amenably embedded G (a term we will define below). Under these conditions, the conclusions of Theorem 1.2 can be significantly strengthened, and we prove a mean and pointwise ergodic theorem for the normalized sampling operators in every L^p, $1 < p < \infty$. Under a suitable strong spectral gap assumption, the pointwise ergodic theorem holds with a rate of convergence, provided the function is subanalytic.

We note that in the generality in which Theorem 1.2 and Theorem 1.3 are stated, the quantitative statement for subanalytic functions is optimal, and therefore the statements of the ergodic theorems are of optimal form.

Our fourth main result, Theorem 1.4, assumes that H is semisimple, that the sets G_t are defined by a norm, and that the volume of the sets H_t is purely exponential. When H acts with a strong spectral gap on G/Γ, we prove mean, maximal and pointwise ergodic theorems with exponentially fast rate of convergence for the normalized sampling operators, for all functions in $L^p(D)$, $1 < p < \infty$. This result is of optimal form, and dispenses entirely with the assumption that the function is subanalytic.

Finally, in Theorem 1.6 we note that the results just stated imply a wide variety of ratio ergodic theorems on homogeneous spaces.

Let us now turn to stating the main results in precise terms.

1.3. **Polynomial normalization of the sampling operators.**

Theorem 1.1. *Assume that*

- *G is an arbitrary almost algebraic group,*
- *for $x \in X$, the stability group $\mathrm{Stab}_G(x) = H$ is of finite index is an almost direct product of a compact subgroup and an abelian diagonalisable subgroup,*
- *the action of Γ on the homogeneous space X is ergodic.*

Then there exist $b \in \mathbb{N}_{>0}$ and $t_0 \in \mathbb{R}_{>0}$ such that the sampling operators

$$\pi_X(\tilde{\lambda}_t)\phi(x) := \frac{1}{t^b} \sum_{\gamma \in \Gamma_t} \phi(x\gamma)$$

satisfy the following :

(i) Strong maximal inequality. *For every $1 < p \le \infty$, compact domain D of X, and $\phi \in L^p(D)$,*

$$\left\| \sup_{t \ge t_0} |\pi_X(\tilde{\lambda}_t)\phi| \right\|_{L^p(D)} \ll_{p,D} \|\phi\|_{L^p(D)}.$$

(ii) Mean ergodic theorem. *For every $1 \le p < \infty$, compact domain D of X, and $\phi \in L^p(D)$,*

$$\left\| \pi_X(\tilde{\lambda}_t)\phi(x) - \int_X \phi \, d\nu \right\|_{L^p(D)} \to 0$$

as $t \to \infty$, where ν is a (nonzero) G-invariant measure on X.

(iii) Pointwise ergodic theorem. *For every $1 \le p \le \infty$, compact domain D of X, and $\phi \in L^p(D)$,*

$$\lim_{t \to \infty} \pi_X(\tilde{\lambda}_t)\phi(x) = \int_X \phi \, d\nu$$

for almost every $x \in X$.

(iv) Quantitative mean ergodic theorem in Sobolev spaces. *Assume, in addition, that the group G is semisimple, Γ is an irreducible lattice in G, and G has a strong spectral gap in $L^2_0(G/\Gamma)$. Then there exists $l \in \mathbb{N}$ such that for every $1 < p < q \le \infty$, compact domain D of X, and $\phi \in L^q_l(D)^+$, the following estimate holds with $\delta_{p,q} > 0$,*

$$\left\| \pi_X(\tilde{\lambda}_t)\phi(x) - \int_X \phi \, d\nu \right\|_{L^p(D)} \ll_{p,q,D} t^{-\delta_{p,q}} \|\phi\|_{L^q_l(D)}$$

for all $t \ge t_0$.

Let us illustrate Theorem 1.1 by giving a pointwise ergodic theorem for an action of a solvable group of exponential growth on a space with infinite measure.

Let Δ be a lattice in $\mathbb{R}^d$, and let a be a $\mathbb{R}$-diagonalisable hyperbolic element of $\mathrm{SL}_d(\mathbb{R})$ that leaves Δ invariant. The group $\Gamma := \langle a \rangle \ltimes \Delta$ is a lattice in $G = \mathbb{R} \ltimes \mathbb{R}^n$, and acts on $\mathbb{R}^d$ by affine transformations:

$$x \cdot (a^n, v) = xa^n + v, \qquad x \in \mathbb{R}^d, \ \ (a^n, v) \in \Gamma. \tag{1.1}$$

Let $\lambda_{\max} > 1$ denote the maximum of absolute values of the eigenvalues of a and $\lambda_{\min} < 1$ denotes the minimum of absolute values of the eigenvalues of a. We fix a norm on $\mathbb{R}^d$ and consider the averaging sets

$$\Gamma_t = \{(a^n, v) \in \Gamma : \ n \in [t/\log(\lambda_{\min}), t/\log(\lambda_{\max})], \ \log \|v\| \le t\} \tag{1.2}$$

Then Theorem 1.1 applies to the averages $\sum_{\gamma\in\Gamma_t}\phi(x\gamma)$ on $\mathbb{R}^d$. In particular, for every $\phi\in L^1(\mathbb{R}^d)$ with compact support,

$$\lim_{t\to\infty}\frac{1}{t}\sum_{\gamma\in\Gamma_t}\phi(v\gamma)=\frac{1}{\mathrm{vol}(\mathbb{R}^d/\Delta)}\int_{\mathbb{R}^d}\phi(x)\,dx\quad\text{for almost every } v\in\mathbb{R}^d. \tag{1.3}$$

Note that while the cardinality of the sets Γ_t grows exponentially (namely, $|\Gamma_t|\sim c\,e^{dt}t$ as $t\to\infty$ with $c>0$), the correct normalisation turns out to be linear in this case. We refer to Section 11.6 below where this example is discussed in detail.

Finally, we note that Theorem 1.1 holds as stated for the sets Γ_t defined when the homogeneous polynomial P is replaced by any vector space norm, on not necessarily a polynomial one (see Remark 8.7 below.)

1.4. **Exponential normalization of the sampling sets.** We now turn to consider the situation where the growth of the sets H_t is exponential, and begin by stating our second main result.

Theorem 1.2. *Assume that*

- *the group G is semisimple, Γ is an irreducible lattice in G, and G has a strong spectral gap in $L^2_0(G/\Gamma)$,*
- *for $x\in X$, the stability group $H=\mathrm{Stab}_G(x)$ is not of finite index in an almost direct product of a compact subgroup and an abelian $\mathbb{R}$-diagonalisable subgroup,*
- *the action of Γ on the homogeneous space $X=G/H$ is ergodic.*

Then there exist $a\in\mathbb{Q}_{>0}$, $b\in\mathbb{N}_{\ge 0}$ and $t_0\in\mathbb{R}_{>0}$ such that the averages

$$\pi_X(\tilde\lambda_t)\phi(x):=\frac{1}{e^{at}t^b}\sum_{\gamma\in\Gamma_t}\phi(x\gamma)$$

satisfy for some $l\in\mathbb{N}_{\ge 0}$,

(i) Strong maximal inequality. *For every $1<p\le\infty$, compact doamin D of X, and $\phi\in L^p_l(D)^+$,*

$$\left\|\sup_{t\ge t_0}|\pi_X(\tilde\lambda_t)\phi|\right\|_{L^p(D)}\ll_{l,p,D}\|\phi\|_{L^p_l(D)}.$$

(ii) Mean ergodic theorem. *There exists a family of absolutely continuous measures $\{\nu_x\}_{x\in X}$ on X, with positive continuous densities such that for every $1\le p<\infty$, compact domain D of X, and $\phi\in L^p(D)$,*

$$\left\|\pi_X(\tilde\lambda_t)\phi(x)-\int_X\phi\,d\nu_x\right\|_{L^p(D)}\to 0$$

as $t\to\infty$.

(iii) Pointwise ergodic theorem. *For every $1<p\le\infty$, compact domain D of X, and bounded $\phi\in L^p_l(D)^+$,*

$$\lim_{t\to\infty}\pi_X(\tilde\lambda_t)\phi(x)=\int_X\phi\,d\nu_x$$

for almost every $x\in X$.

(iv) Quantitative pointwise ergodic theorem. *For every $1 < p \le \infty$, compact domain D of X, and a nonnegative continuous subanalytic function $\phi \in L_l^p(D)$, the following asymptotic expansion holds*

$$\pi_X(\tilde{\lambda}_t)\phi(x) = \int_X \phi\, d\nu_x + \sum_{i=1}^{b} c_i(\phi, x)t^{-i} + O_{x,\phi}\left(e^{-\delta(x,\phi)t}\right)$$

for almost every $x \in X$ and all $t \ge t_0$ with some $\delta(x, \phi) > 0$.

Let us give an example of application of Theorem 1.2. Let Γ be a lattice in $\mathrm{SL}_d(\mathbb{R})$ and $\Gamma_t = \{\gamma \in \Gamma : \log \|\gamma\| \le t\}$ denote the norm balls with respect to the standard Euclidean norm $\|\gamma\| = \left(\sum_{i,j=1}^d \gamma_{ij}^2\right)^{1/2}$. We consider the action of Γ on the projective space $\mathbb{P}^{d-1}(\mathbb{R})$. Then Theorem 1.2(iv) implies that for any nonnegative continuous subanalytic function $\phi \in L_l^p(\mathbb{P}^{d-1}(\mathbb{R}))$ with $p > 1$ (for some explicit $l \ge 0$) and for almost every $v \in X$, there exists $\delta > 0$ such that

$$\frac{1}{e^{(d^2-d)t}} \sum_{\gamma \in \Gamma_t} \phi(v\gamma) = c_d(\Gamma) \int_{\mathbb{P}^{d-1}(\mathbb{R})} \phi(w)\, d\xi(w) + O_{\phi,v}(e^{-\delta t}), \tag{1.4}$$

where $c_d(\Gamma) > 0$, $\delta = \delta(v, \phi) > 0$, and ξ denote the $\mathrm{SO}_d(\mathbb{R})$-invariant probability measure on $\mathbb{P}^{d-1}(\mathbb{R})$. This example is in more detail discussed in Section 11.2. Further examples and applications of Theorem 1.2 are discussed in Section 11.

Note that in Theorem 1.2 the group H can be a solvable, for example. In that case, no rate of convergence can possibly hold for the operators $\pi_X(\tilde{\lambda}_t)$ acting in Lebesgue space, and results in Sobolev spaces are the best that can be achieved. The same remark applies of course to Theorem 1.1.

For a special class of homogeneous spaces X, we obtain an improved version of Theorem 1.2 with L^p-norms in place of Sobolev norms, to which we now turn.

1.5. **Non-amenable stabilizers and quantitative ergodic theorems.**

Theorem 1.3. *Assume that at least one of the following conditions is satisfied.*

- *G is an arbitrary almost algebraic group, the stability group $H = \mathrm{Stab}_G(x)$ is semisimple, and H has a strong spectral gap in $L_0^2(G/\Gamma)$,*
- *G is a semisimple group which has a strong spectral gap in $L_0^2(G/\Gamma)$, and H is any almost algebraic subgroup which is unimodular and non-amenably embedded (see Definition 9.8 below).*

Then there exist $a \in \mathbb{Q}_{>0}$, $b \in \mathbb{N}_{\ge 0}$ and $t_0 \in \mathbb{R}$ such that the normalized sampling operators

$$\pi_X(\tilde{\lambda}_t)\phi(x) := \frac{1}{e^{at}t^b} \sum_{\gamma \in \Gamma_t} \phi(x\gamma)$$

satisfy the following

(i) Strong maximal inequality. *For every $1 < p \le \infty$, compact doamin D of X, and $\phi \in L^p(D)$,*

$$\left\| \sup_{t \ge t_0} |\pi_X(\tilde{\lambda}_t)\phi| \right\|_{L^p(D)} \ll_{p,D} \|\phi\|_{L^p(D)}.$$

(ii) Pointwise ergodic theorem. *For every $1 < p \le \infty$, compact domain D of X, and $\phi \in L^p(D)$,*

$$\lim_{t\to\infty} \pi_X(\tilde{\lambda}_t)\phi(x) = \int_X \phi \, d\nu_x$$

for almost every $x \in X$.

(iii) Quantitative pointwise ergodic theorem. *For every $1 < p \le \infty$, compact domain D of X, and a nonnegative continuous subanalytic function ϕ with* $\mathrm{supp}(\phi) \subset D$, *the following asymptotic expansion holds*

$$\pi_X(\tilde{\lambda}_t)\phi(x) = \int_X \phi \, d\nu_x + \sum_{i=1}^{b} c_i(\phi, x) t^{-i} + O_{\phi,x}\left(e^{-\delta(x,\phi)t}\right)$$

for almost every $x \in X$ and all $t \ge t_0$ with some $\delta(x,\phi) > 0$.

To exemplify our general results, let us consider the action of a lattice Γ in the orthogonal group $\mathrm{SO}_{d,1}(\mathbb{R})^0$ on the quadratic surface

$$X = \{x \in \mathbb{R}^{d+1} : x_1^2 + \cdots + x_d^2 - x_{d+1}^2 = 1\},$$

The space X is known as the de-Sitter space, and the problem of distribution of orbits of Γ in X was raised by Arnol'd (see [Ar, 1996-15, 2002-16]). Theorem 1.1 solves the problem for $d = 2$, and Theorem 1.3 solves this problem for general $d \ge 3$. It will be convenient to use the polar coordinate system on X:

$$\mathbb{R} \times S^{d-1} \to X : (r, \omega) \mapsto (\omega_1 \cosh r, \ldots, \omega_d \cosh r, \sinh r). \tag{1.5}$$

For $d = 2$, we obtain from Theorem 1.1 that for every $\phi \in L^1(X)$ with compact support and for almost every $v \in X$,

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\gamma\in\Gamma_t} \phi(v\gamma) = c_2(\Gamma) \int_X \phi(r,\omega)\,(\cosh r)\, dr\, d\omega. \tag{1.6}$$

for some $c_2(\Gamma) > 0$. For $d \ge 3$, we obtain from Theorem 1.3 that for every nonnegative continuous subanalytic function ϕ with compact support and almost every $v \in X$, the following asymptotic expansion holds

$$\frac{1}{e^{(d-2)t}} \sum_{\gamma\in\Gamma_t} \phi(v\gamma) = \frac{c_d(\Gamma)}{(1+v_d^2)^{(d-2)/2}} \int_X \phi(r,\omega)\, \frac{(\cosh r)^{d-1} dr\, d\omega}{(1+(\sinh r)^2)^{(d-2)/2}} + O_{v,\phi}(e^{-\delta t}) \tag{1.7}$$

for some $c_d(\Gamma) > 0$ and $\delta = \delta(v,\phi) > 0$. Note that the limit measure in this case is not a Γ-invariant measure, and moreover it depends nontrivially on the initial point v. Further applications of Theorem 1.3 are discussed in Section 11.

To motivate the discussion immediately below, let us note that in the present example it is in fact possible to obtain a much stronger conclusion. Both the restriction that ϕ is subanalytic, as well as the dependence $\delta(v,\phi)$ on v and ϕ can be dispensed with, as follows from Theorem 1.4. We will discuss this example in more detail in Section 11.1 below.

1.6. **Volume regularity and quantitative ergodic and ratio theorems.**

1.6.1. *On the role of volume regularity in the proofs.* A interesting feature exhibited in Theorem 1.2(iv) and Theorem 1.3(iii) is that the quality of quantitative ergodic theorems stated in them is genuinely constrained. While we assume spectral gap conditions which imply exponential norm decay of the averaging operators supported on H_t acting on $L^2_0(G/\Gamma)$, the dual operators $\pi_X(\tilde{\lambda}_t)$ on $L^2(D)$ do not satisfy such an exponential norm or pointwise decay estimate. For subanalytic function on D, the asymptotic expansion :

$$\pi_X(\tilde{\lambda}_t)\phi(x) = \int_X \phi\, d\nu_x + \sum_{i=1}^{b} c_i(\phi, x)t^{-i} + O_{\phi,x}\left(e^{-\delta(x,\phi)t}\right)$$

which holds for almost every $x \in X$ with some $\delta(x, \phi) > 0$, implies that if $b > 0$ the rate of almost sure convergence to the limiting distribution is at least t^{-1}, but typically not faster, so it is *not* exponential. Let us now explain the reason for the occurrence of this phenomenon, and then state a substantial improvement to the quantitative ergodic theorems under suitable conditions.

A fundamental reduction that appears repeatedly in our analysis below is the comparison of the normalized sampling operators on a Γ-orbit in X :

$$\pi_X(\lambda_t)\phi(x) = \frac{1}{\mathrm{vol}(H_t)} \sum_{\gamma\in\Gamma_t} \phi(x\gamma),$$

with the normalized sampling operators on the G-orbit in X :

$$\pi_X(\lambda^G_t)\phi(x) = \frac{1}{\mathrm{vol}(H_t)} \int_{g\in G_t} \phi(xg)dm(g)$$

where here we use the intrinsic normalization by $\mathrm{vol}(H_t)$ rather than by $e^{at}t^b$.

It will develop that the difference $\pi_X(\lambda_t) - \pi_X(\lambda^G_t)$ can be estimated very well under very general assumptions. It converges to zero in the mean and pointwise, and in fact with an exponentially fast quantitative rate if the corresponding operators supported on H_t satisfy the spectral gap estimates and quantitative ergodic theorem in $L^2(G/\Gamma)$. Furthermore, $\pi_X(\lambda^G_t)\phi(x) - \int_D \phi\, d\nu_x$, converges to zero almost surely, and this identifies the limit of $\pi_X(\lambda_t)\phi(x)$ in the mean and pointwise ergodic theorems, as the limiting density ν_x. To obtain a quantitative ergodic theorem for $\pi_X(\lambda_t)$ all we need to do is make the latter convergence result quantitative. To understand the limitations here, let us note the following alternative expression for $\pi_X(\lambda^G_t)$ and its connection with the limiting distribution $\tilde{\nu}_x$. Let p_X denotes the canonical projection from G to $X = H \setminus G$, let s denotes a measurable section from X to G which is bounded on compact sets, and ξ denote the canonical density on X. Then by the discussion preceding equation (4.5) below, with

$H_t[g_1, g_2] = H \cap g_1 G_t g_2^{-1}$,

$$\begin{aligned}\pi_X(\lambda_t^G)\phi(x) &= \frac{1}{\rho(H_t)}\int_{G_t}\phi(\mathsf{p}_X(\mathsf{s}(x)g))\,dm(g)\\ &= \frac{1}{\rho(H_t)}\int_{(y,h):\,\mathsf{s}(x)^{-1}h\mathsf{s}(y)\in G_t}\phi(\mathsf{p}_X(h\mathsf{s}(y)))\,d\rho(h)d\xi(y)\\ &= \int_D \phi(y)\frac{\rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])}{\rho(H_t)}\,d\xi(y).\end{aligned}$$

Now for every $g_1, g_2 \in G$, we shall show that the limit

$$\Theta(g_1, g_2) := \lim_{t\to\infty}\frac{\rho(H_t[g_1, g_2])}{\rho(H_t)}$$

exists (see Lemma 8.5), and defines the family of limiting distributions ν_x , $x \in X$ (see (2.7)).

Consequently, the convergence properties of $\pi_X(\lambda_t^G)\phi(x)$ to $\int_D \phi(y)d\nu_x(y)$ are determined by the convergence properties of $\rho(H_t[\mathsf{s}(x), \mathsf{s}(y)])/\rho(H_t)$ to the limiting distribution ν_x. Thus the regularity properties of the volume of the sets H_t are crucial, and control the final conclusion in the quantitative ergodic theorem. The parameter a and b that appear in the quantitative results for the operators $\pi_X(\lambda_t)$ acting on subanalytic functions in Theorem 1.2(iv) and Theorem 1.3(iii), are those that appear in the development $\mathrm{vol}(H_t) = e^{at}(c_b t^b + \cdots + c_0) + O(e^{(a-\delta_0)t})$ given by Theorem 8.1(i) below.

The foregoing discussion shows that the results for subanalytic functions are optimal as stated, but suggest that under the stronger assumption that the asymptotic development of the volume holds with $b = 0$, the results in the ergodic theorem should be stronger. We will now state our fourth main result, which gives the optimal formulation of quantitative ergodic theorems under this additional assumption. As we will see below, this assumption is satisfied by a large collection of important examples.

Theorem 1.4. *Assume that the following conditions are satisfied.*

- $G \subset \mathrm{SL}_d(\mathbb{R})$ *is an arbitrary almost algebraic group,*
- *the stability group* $H = \mathrm{Stab}_G(x) \subset G$ *is semisimple and has a strong spectral gap in* $L_0^2(G/\Gamma)$*,*
- *the homogeneous polynomial* P *is replaced by a norm on* $\mathrm{Mat}_d(\mathbb{R})$*,*
- *the volumes of* H_t *satisfy* $\mathrm{vol}(H_t) \sim c\,e^{at}$ *as* $t \to \infty$*, with* $a, c > 0$*.*

Then the normalized sampling operators $\pi_X(\tilde{\lambda}_t)\phi(x) = \frac{1}{e^{at}}\sum_{\gamma\in\Gamma_t}\phi(x\gamma)$ *satisfy, for every compact domain* $D \subset X$ *and* $t \geq t_0$*,*

(i) Quantitative mean ergodic theorem in Lebesgue spaces. *For* $1 < p < \infty$*, a suitable* $\delta_p > 0$ *independent of* D*, and for every* $\phi \in L^p(D)$

$$\left\|\pi_X(\tilde{\lambda}_t)\phi(x) - \int_X \phi\,d\nu_x\right\|_{L^p(D)} \ll_{p,D} \|\phi\|_{L^p(D)}\,e^{-\delta_p t}\,.$$

(ii) Quantitative maximal ergodic theorem in Lebesgue spaces. *The family* $\pi_X(\tilde{\lambda}_t)$ *satisfies the* (L^p, L^w)*-exponential strong maximal inequality (for some*

$1 < w < p < \infty$), namely there exists $\delta_{p,w} > 0$ such that for every $\phi \in L^p(D)$,

$$\left\|\sup_{t\ge t_0} e^{\delta_{p,w}t}\left|\pi_X(\tilde\lambda_t)\phi(x) - \int_D \phi\, d\nu_x\right|\right\|_{L^w(D)} \ll_{p,w,D} \|\phi\|_{L^p(D)}.$$

(iii) Uniform quantitative pointwise ergodic theorem in Lebesgue spaces. *For every $\phi \in L^p(D)$, $1 < p < \infty$, and almost every $x \in X$,*

$$\left|\pi_X(\tilde\lambda_t)\phi(x) - \int_X \phi\, d\nu_x\right| \ll_{x,\phi,D} e^{-\delta_p t},$$

where $\delta_p > 0$ is independent of x, ϕ and D.

Remark 1.5. (1) We note that when the volume growth is not purely exponential, namely when $b > 0$, the quantitative mean and pointwise ergodic theorems hold as stated in Theorem 1.4(i) and (iii), but the speed is $t^{-\eta_p}$(with $\eta_p > 0$) rather than $e^{-\delta_p t}$ (see Subsection 10.4).

(2) In [GN3, Thm. 1.6], we have applied Theorem 1.4 to the case of a dense S-arithmetic lattice in a connected semisimple Lie group, which acts by isometries on the group variety itself. We have established there an exponentially fast quantitative equidistribution theorem, *for every starting point, and with a fixed rate*, in the case of Hölder functions. The statement of [GN3, Thm. 1.6] refers to the case of norms with purely exponential growth of balls. When the growth is not purely exponential, the rate of equidistribution is polynomial.

We note that the assumptions of Theorem 1.4 are verified in many interesting cases, when we replace the polynomial P by a norm with purely exponential growth of balls. These include Examples 11.1 of the action of $\mathrm{SO}_{n,1}(\mathbb{R})^0$ on de-Sitter space, Example 11.4 of dense subgroups of semisimple Lie groups acting by translation on the group, Example 11.5 of distribution of values of indefinite quadratic forms, and Example 1.7 of affine actions of lattices. These examples will be explained in detail in §11.

1.6.2. *Ratio ergodic theorems.* In the case of a single transformation acting on an infinite, σ-finite measure space, it is well-known that a general ratio ergodic theorem holds, and it is natural to consider ergodic ratio theorems in our context as well. The results stated above describe the limiting behavior of the operators $\pi_X(\tilde\lambda_t)$, and so it immediately follows that they imply a limit theorem for their ratios. We record this fact in the following result.

Theorem 1.6. *Let G, H, X, Γ and P be as in the previous section. Consider a compact domain $D \subset X$, and any two functions $\phi, \psi \in L^p(D)$ of compact support contained in D, with ψ non-negative and non-zero.*

(i) Ratio ergodic theorem. *Assume that the conditions of Theorem 1.1, or Theorem 1.3 are satisfied. Then as $t \to \infty$,*

$$\frac{\sum_{\gamma\in\Gamma_t}\phi(x\gamma)}{\sum_{\gamma\in\Gamma_t}\psi(x\gamma)} \longrightarrow \frac{\int_D \phi\, d\nu_x}{\int_D \psi\, d\nu_x}$$

for almost every $x \in D$. Under the conditions of Theorem 1.2, the same result applies provided ϕ and ψ as above are in $L^p_l(D)$.

(ii) Quantitative ratio ergodic theorem. *Assume that the conditions of Theorems 1.2 (iv) or Theorem 1.3(iii) hold. Let b and δ be the volume growth parameters stated there. Assume that $\phi, \psi \in C_c(D)$ are subanalytic and non-negative (and in $L^p_l(D)$ when assuming the conditions of Theorem 1.2(iv)). Then convergence in the ratio ergodic theorem takes place at the rate*

$$\left| \frac{\sum_{\gamma\in\Gamma_t} \phi(x\gamma)}{\sum_{\gamma\in\Gamma_t} \psi(x\gamma)} - \frac{\int_D \phi \, d\nu_x}{\int_D \psi \, d\nu_x} \right| \ll_{x,\phi,\psi} E(t)$$

where $E(t) = \frac{1}{t}$ if $b > 0$, and $E(t) = e^{-\delta_p t}$ for some $\delta_p(x, \phi, \psi) > 0$, if $b = 0$.

(iii) Uniform quantitative ratio ergodic theorem. *Under the conditions of Theorem 1.4, the exponential rate of convergence δ_p in the ratio ergodic theorem stated in (ii) is independent of ϕ, ψ, x and D.*

Consider now the problem of establishing equidistribution for the ratio of averages, namely convergence for every single starting point, when the functions are continuous, or satisfy (say) a Hölder regularity condition. In the case of dense subgroups of isometries of Euclidean spaces this problem was introduced half a century ago by Kazhdan [K]. In [GN3], we have established a quantitative ratio equidistribution theorem for certain dense subgroups of semisimple Lie groups acting on the group manifold, with respect to Hölder functions. This result has the remarkable feature that it holds for every single starting point $x \in X$, with the same rate of exponentially fast convergence (provided the volume growth of balls is purely exponential). This result is based on the quantitative ergodic theorem in L^2 stated in Theorem 1.4.

1.7. **Comments on the development of the duality principle.** Given an lcsc group G, and two closed subgroups H_1 and H_2, the principle of duality (namely, that properties of the H_1-action on G/H_1 and the properties of the H_2-actions on $H_1 \setminus G$ are closely related) has been a mainstay of homogeneous dynamics for at least half a century. It has been used in a diverse array of different applications, as demonstrated in the list below. Needless to say, the list constitutes just a small sample of the extensive literature in homogeneous dynamics which can be construed as applying duality arguments in one form or another.

As to general duality properties, let us mention the following often-used fundamental facts.

- Minimality : H_1 is minimal on G/H_2 if and only if H_2 is minimal on $H_1 \setminus G$. This fact and its application to Diophantine approximation in the context of flows on nilmanifolds has been discussed already in [AGH], see e.g. Chapter VIII by Auslander and Green and Chapter XI by Greenberg.
- Ergodicity : H_1 is ergodic on G/H_2 if and only if H_2 is ergodic on $H_1 \setminus G$. This result is known as Moore's ergodicity theorem [Mo].
- Amenability : H_1 acts amenably on G/H_2 if and only if H_2 acts amenably on $H_1 \setminus G$. The notion of amenability used here is that defined by Zimmer, see [Z1].

Let us now restrict the discussion to the case when G is semisimple, $H_2 = \Gamma$ is an irreducible lattice, and H_1 is a minimal parabolic subgroup, or a unipotent subgroup of it.

- Unique ergodicity of horocycle flows U on G/Γ when $G = \mathrm{SL}_2(\mathbb{R})$ has been proved by Furstenberg [F] using the dynamical properties of the Γ-action on $U \setminus G$. This method has subsequently been generalized by Veech [V2] for semisimple groups.
- Dani and Raghavan [DR] studied orbits of frames under discrete linear groups, and in particular considered the connection between the density properties of Γ acting on $U \setminus G$, and U acting on G/Γ, where U is a horospherical unipotent subgroup. Minimality of the action of a horospherical group when the lattice is uniform has been proved earlier by Veech [V1]. Related equidistribution results (namely, the convergence of sampling operators $\pi_X(\lambda_t)$ in the space of continuous functions) were established in [Le1, No, G2].
- Dani [D1] established the topological version of Margulis factor theorem, namely that continuous Γ-equivariant factors of the boundary $P \setminus G$ are necessarily of the form $Q \setminus G$, where Q is a parabolic subgroup containing P. The proof uses the minimality properties of the P-action on G/Γ. Special cases have previously been established by Zimmer [Z2], Spatzier [Sp], and a generalization was established by Shah [Sh].
- Using duality-type arguments, equidistribution of the lattice orbits on the boundary $P\backslash G$ for $G = \mathrm{SL}_d(\mathbb{R})$ was established in [G1]. The method utilizes the equidistribution of averages on P acting on $C_c(G/\Gamma)$. Subsequently equidistribution of the Γ-orbits on the boundary was established for general semisimple groups in [GO], and using a different method in [GM].

Let us now proceed with G a semisimple group, $H_2 = \Gamma$ an irreducible lattice subgroup, $H_1 = H$ an algebraic subgrpup, and assume that $H \setminus G$ has a G-invariant measure.

- The problem of counting the number of points of an orbit in a ball for isometry groups of manifolds of (variable) negative curvature was already considered in Margulis' thesis (see [Ma1]). The basic connection between this problem and the mixing property of the geodesic flow established there has been a major source of influence in the development of duality arguments.
- Several problems in Diophantine approximation in Euclidean spaces can be approached via duality arguments, a fact originally due to Dani [D2] who studied it systematically. This fact is referred to as the Dani correspondence by Kleinbock and Margulis [KM1, KM2], who have extended it further.
- Duality considerations also played an important role in Margulis' celebrated solution of Oppenheim conjecture (see [Ma2]), where the method of proof utilizes the dynamics of $\mathrm{SO}_{2,1}(\mathbb{R})$ in the space of lattices $\mathrm{SL}_3(\mathbb{R})/\mathrm{SL}_3(\mathbb{Z})$ in order to analyse the values of the quadratic form on integer points, namely on an orbit of the lattice $\mathrm{SL}_3(\mathbb{Z})$. A quantitative approach to this problem

was subsequently developed by Dani and Margulis [DM] and Eskin, Margulis and Mozes [EMM].

- The general lattice point counting problem on homogeneous spaces is to give precise asymptotics for the number of points in discrete orbits of Γ in $H \setminus G$ contained in a ball, and the most studied case is when H is a symmetric subgroup. By duality, any such discrete orbit determines canonically a closed orbit of H in G/Γ. Quantitative results in the lattice point counting problem can be established by applying harmonic analysis to the "H-periods" of suitable automorphic functions on G/Γ, namely by estimating their integral on such closed orbits. This spectral approach in the case of higher rank semisimple groups and symmetric subgroups was first applied by Duke, Rudnick and Sarnak [DRS].
- Establishing the main term in the lattice point counting problem for Γ on $H \setminus G$ can be reduced to establishing equidistribution of translates of the closed H-orbit $H\Gamma \subset G/\Gamma$. Eskin and McMullen [EM] have established the equidistribution of these translates using the mixing property on A on G/Γ. Eskin, Mozes and Shah [EMS] have established equidistribution of translates of a closed H-orbit using the theory of unipotent flows on G/Γ. The mixing method can also be made quantitative, as shown by Maucourant [Mau] and [BO], using estimates for rates of mixing in Sobolev spaces.
- An important development regarding the duality principle and equidistribution of lattice actions on homogeneous varieties is due to Ledrappier. In [Le1, Le2] he considered the action of a lattice subgroup of $\mathrm{SL}_2(\mathbb{R})$ on the plane $\mathbb{R}^2$, and used duality arguments to establish convergence of the sampling operators supported on Γ_t, in the space of continuous functions. Ledrappier's results revealed for the first time several of the remarkable features that arise in the context of infinite-volume homogeneous spaces, including the appearance of a limiting density different than the invariant measure, the fact that the limiting density is not necessarily invariant under the lattice action, and the fact that it depends on the initial point in the orbit under consideration. Subsequently Ledrappier and Pollicott [LP][LP2] have generalised these results to SL_2 over other fields. Another appearance of duality argument is due to Maucourant, who in his thesis considered the Γ-action on G/A where A is an $\mathbb{R}$-split torus, via ergodic theorems for A acting on G/Γ. A systematic general approach to the duality principle for a large class of groups and homogeneous spaces, including an analysis of the limiting distribution and the requisite properties of volume regularity was carried out in [GW], and applied to obtain diverse equidistribution results for actions of lattices on homogeneous spaces.
- The quantitative solution of the lattice point counting problem which appears in [GN2] can be viewed as a quantitative duality argument, applied to the case where the subgroup H is in fact equal to G.
- The Boltzmann-Grad limit for periodic Lorentz gas has been studied by Marklof and Strömbergsson [MS1, MS2]. Using periodicity, one can reduce the original problem to analysing distribution on the space of lattices $\mathrm{SL}_d(\mathbb{R})/\mathrm{SL}_d(\mathbb{Z})$, which can be treated using the theory of unipotent flows.

Given the centrality of the duality principle in homogeneous dynamics, it is a most natural problem to establish mean and pointwise ergodic theorems in Lebesgue spaces for lattice actions on homogeneous varieties, including a rate of convergence in the presence of a spectral gap. The present paper is devoted to the solution of this problem, which we expect will have several significant applications, including a quantitative approach to Diophantine approximation on homogeneous algebraic varieties, developing [GGN] further. However, we are not aware of any previous results in the literature establishing ergodic theorems in Lebesgue spaces for infinite volume homogeneous spaces.

1.8. **Organization of the paper.** In Section 2 we set-up notations that will be used throughout the paper. Sections 3–7 form the core of the paper. In this part we develop the general ergodic-theoretic duality framework in the setting of locally compact topological groups and their homogeneous spaces. Section 8 is devoted to the crucial issue of volume regularity, and contains a number of results concerning regularity properties of balls defined by polynomials in algebraic group. In Section 9 we develop ergodic theory of actions of algebraic groups on probability measure space, which are subsequently used as an input for the general duality principle developed in the previous sections. In Section 10 we complete the proof of the four main theorems stated in the introduction. Finally, Section 11 is devoted to examples illustrating our main results.

Acknowledgements. The authors would like to thank Barak Weiss for his very valuable contributions to this project. We also would like to express our gratitude to the Ergodic Theory Group at the Fédération Denis Poisson who gave us an opportunity to explain the present work in the lecture series “Théorie ergodique des actions de groupe” held in Tours in April, 2011. A.G. would like to thank the Princeton University and especially E. Lindenstrauss for their hospitality during 2007–2008 academic year when part of this work has been completed.

2. Basic notation

In this section we introduce notation that will be used throughout the paper.

2.1. **Sobolev norms.** Given a compact domain D of $\mathbb{R}^d$, we denote by $L^p_l(D)$ the space of functions ϕ on $\mathbb{R}^d$ with support contained in D for which all distributional derivatives of ϕ of order at most l are in $L^p(\mathbb{R}^d)$. The space $L^p_l(D)$ is equipped with the Sobolev norm

$$\|\phi\|_{L^p_l(D)} = \sum_{0\le |s|\le l} \|\mathcal{D}^s\phi\|_{L^p(\mathbb{R}^d)},$$

where $\mathcal{D}^s$ denotes the partial derivative corresponding to the multi-index s. More generally, let D be a compact subset of a manifold X of dimension d. We pick a system of coordinate charts $\{\omega_i : \mathbb{R}^d \to M\}_{i=1}^n$ that cover D and a partition of unity $\{\psi_i\}_{i=1}^n$ subordinate to this system of charts such that $\sum_{i=1}^n \psi_i = 1$ on D. We denote by $L^p_l(D)$ the space of functions ϕ on X with support contained in D such that for every $i = 1, \dots, n$, the function $(\phi \cdot \psi_i) \circ \omega_i$ belongs to $L^p_l(\mathbb{R}^d)$. The space

$L_l^p(D)$ is equipped with the Sobolev norm

$$\|\phi\|_{L_l^p(D)} = \sum_{i=1}^{n} \|(\phi \cdot \psi_i) \circ \omega_i\|_{L_l^p}. \tag{2.1}$$

Clearly, this definition of the Sobolev norm depends on a choice of the system of coordinate charts and the partition of unity, but using compactness of D, one checks that different choices lead to equivalent norms, and for this reason we have suppressed ω_i's and ψ_i's from the notation.

2.2. **Homogeneous spaces and measures.** Let G be a locally compact second countable (lcsc) group, H a closed subgroup of G, and $X := H\backslash G$ is a homogeneous space of G. In the case when our discussion involves Sobolev norms, we always assume, in addition, that G is a connected Lie group and H is a closed Lie subgroup. Then X has the structure of a smooth manifold. We denote by p_X the natural projection map

$$\mathsf{p}_X : G \to X : g \mapsto Hg,$$

and by $\mathsf{s} : X \to G$ a measurable section of this map, i.e., a map such that $\mathsf{p}_X \circ \mathsf{s} = id$. Since G is locally compact, such a section always exists, with the additional property that it is bounded on compact sets. We also use the notation:

$$\mathsf{r} := \mathsf{s} \circ \mathsf{p}_X : G \to G \quad \text{and} \quad \mathsf{h}(g) := \mathsf{r}(g)g^{-1} \in H.$$

Then for $g \in G$,

$$g = \mathsf{h}(g)^{-1}\mathsf{r}(g). \tag{2.2}$$

We note that these maps satisfy

$$\mathsf{p}_X(hg) = \mathsf{p}_X(g) \quad \text{and} \quad \mathsf{h}(hg) = \mathsf{h}(g)h^{-1} \tag{2.3}$$

for $g \in G$ and $h \in H$.

Let Γ be a lattice subgroup in G, that is, a discrete subgroup of finite covolume. We set $Y := G/\Gamma$ and denote by

$$\mathsf{p}_Y : G \to Y : g \mapsto g\Gamma$$

the natural projection map.

Let m be a left Haar measure on G and ρ a left Haar measure H. It follows from invariance that the measure m has the decomposition

$$\int_G f(g)\, dm(g) = \int_{H\times X} f(h \cdot \mathsf{s}(x))\, d\rho(h) d\xi(x), \quad \phi \in L^1(G), \tag{2.4}$$

where ξ is a Borel measure on X. We note that the measure ξ does depend on the section s. If fact, if s_1 and s_2 are two sections, then $\mathsf{s}_1(x)\mathsf{s}_2(x)^{-1} \in H$ and the corresponding measures ξ_1 and ξ_2 on X are related by

$$d\xi_1(x) = \Delta_H(\mathsf{s}_1(x)\mathsf{s}_2(x)^{-1})\, d\xi_2(x), \tag{2.5}$$

where Δ_H denotes the modular function for the Haar measure on H. In particular, when H is unimodular, the measure ξ is canonically defined.

Our arguments below involve functions supported in a compact domain D of X, and the measurable section $\mathsf{s} : D \to G$ which we choose so that $\mathsf{s}(D)$ is bounded in

G. It follows that $\xi(D) < \infty$. We denote by $L^p(D)$ the space of functions ϕ on X with $\mathrm{supp}(\phi) \subset D$ and

$$\|\phi\|_{L^p(D)} := \left(\int_D |\phi|^p \, d\xi \right)^{1/p} < \infty.$$

While this norm depends on a choice of s, it follows from (2.5) that different choices lead to equivalent norms.

Let μ be the Haar measure on $Y = G/\Gamma$ induced by m. Namely, μ is defined by $\mu(A) = m(\mathsf{p}_Y^{-1}(A) \cap \mathcal{F})$ where $\mathcal{F} \subset G$ is a fundamental domain for the right action of Γ on G. We normalise Haar measure m on G so that $\mu(G/\Gamma) = 1$.

Throughout the paper, unless stated otherwise, we use the measure m on G, the measure ξ on X, and the measure μ on Y.

Let $\{G_t\}$ be an increasing sequence of compact subsets of G. For a subset S of G, we set

$$S_t := S \cap G_t,$$

and more generally for $g_1, g_2 \in G$, we set

$$S_t[g_1, g_2] := S \cap g_1 G_t g_2^{-1}. \tag{2.6}$$

Let now H be a closed subgroup. Assuming that the limit

$$\Theta(g_1, g_2) := \lim_{t \to \infty} \frac{\rho(H_t[g_1, g_2])}{\rho(H_t)}$$

exists (see, for instance, Lemma 8.5 below), we consider the family of measures ν_x on X indexed by $x \in X$ and defined by

$$d\nu_x(z) := \Theta(\mathsf{s}(x), \mathsf{s}(z)) \, d\xi(z), \quad z \in X. \tag{2.7}$$

We note that while the measure ξ depends on a choice of the section s, the measures ν_x are canonically defined.

In the case when G is a Lie group and H is a closed subgroup, the measures m and ρ are defined by smooth differential forms on X. Since the section s can be chosen locally smooth, it follows that the measures ν_x are absolutely continuous.

We note that the "skew balls" $H_t(g_1, g_2)$ and the limiting measures ν_x were originally introduced in [GW], in which further details can be found.

2.3. **Ergodic-theoretic terminology.** Let (X, ξ) be a standard Borel space equipped with an action of an lcsc group L which is denoted by π_X. For a family of finite Borel measures ϑ_t on L, we denote by $\pi_X(\vartheta_t)$ the averaging operator defined on measurable functions ϕ on X by

$$\pi_X(\vartheta_t)\phi(x) = \int_L \phi(x \cdot l) \, d\vartheta_t(l), \quad x \in X. \tag{2.8}$$

Definition 2.1. Let $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$ be a Banach space of measurable functions on X.

- We say that the family of measures ϑ_t satisfies the *strong maximal inequality* in $\mathbb{B}$ with respect to a seminorm $\|\cdot\|$ if for every $\phi \in \mathbb{B}$,

$$\left\| \sup_{t \geq 1} |\pi_X(\vartheta_t)\phi| \right\| \ll \|\phi\|_{\mathbb{B}}.$$

- We say that the family of measures ϑ_t satisfies the *mean ergodic theorem* in $\mathbb{B}$ if for every $\phi \in \mathbb{B}$, the sequence $\pi_X(\vartheta_t)\phi$ converges in the norm of $\mathbb{B}$ as $t \to \infty$.
- We say that the family of measures ϑ_t satisfies the *quantitative mean ergodic theorem* in $\mathbb{B}$ with respect to a seminorm $\|\cdot\|$ on $\mathbb{B}$ if for every $\phi \in \mathbb{B}$, the limit of $\pi_X(\vartheta_t)\phi$ as $t \to \infty$ exists in the norm of $\mathbb{B}$ and
$$\left\| \pi_X(\vartheta_t)\phi - \left(\lim_{t\to\infty} \pi_X(\vartheta_t)\phi\right) \right\| \leq E(t)\, \|\phi\|_{\mathbb{B}}$$
where $E(t) \to 0$ as $t \to \infty$.
- We say that the family of measures ϑ_t satisfies the *pointwise ergodic theorem* in $\mathbb{B}$ if for every $\phi \in \mathbb{B}$, the sequence $\pi_X(\vartheta_t)\phi(x)$ converges as $t \to \infty$ for almost every $x \in X$.

We note that our measures θ_t will be either atomic measures on the finite sets Γ_t, or absolutely continuous bounded Borel measures on compact subsets of the lcsc group H. Therefore measurability of the maximal functions when the supremum is taken over all $t \in \mathbb{R}_+$ follows from standard arguments.

The space of particular interest to us is the spaces $L^p(X)$ consisting of L^p-integrable functions on X. However, when (X, ξ) has infinite measure, it is natural to consider a filtration of X by domains D with finite measure and the spaces $L^p(D)$ of L^p-integrable functions with support contained in D. Moreover, when X has a manifold structure, we also consider the spaces $L^p_l(D)$ of Sobolev functions introduced in Section 2.1.

Given a space of functions $\mathbb{B}$, we denote $\mathbb{B}^+$ the cone in $\mathbb{B}$ consisting of nonnegative functions.

2.4. **Orbit sampling operators.** We conclude this section by introducing the operators that will be in the centre of our discussion. Here and throughout the paper we use the notation introduced in Section 2.2.

We consider the normalized orbit sampling operators, defined for a function $\phi : X \to \mathbb{R}$ and $x \in X = H \setminus G$ by
$$\pi_X(\lambda_t)(\phi)(x) := \frac{1}{\rho(H_t)} \sum_{\gamma\in\Gamma_t} \phi(x \cdot \gamma). \tag{2.9}$$
As noted in the discussion of the duality principle in the Introduction, our plan is to deduce the asymptotic behavior of these sampling operators from the asymptotic behavior of the averaging operators on $Y = G/\Gamma$ defined by
$$\pi_Y(\beta_t^{g_1,g_2})(F)(y) := \frac{1}{\rho(H_t[g_1,g_2])} \int_{H_t[g_1,g_2]} F(h^{-1}\cdot y)\, d\rho(h). \tag{2.10}$$
for $g_1, g_2 \in G$, a function $F : Y \to \mathbb{R}$, and $y \in Y$.

To simplify notation, we also set $\beta_t = \beta_t^{e,e}$.

3. The basic norm bounds and the strong maximal inequality

3.1. **Coarse admissibility.** The goal of this section is prove the strong maximal inequality for the operators $\pi_X(\lambda_t)$ defined in (2.9), based on the validity of a maximal inequality for the operators defined in (2.10). We shall use notation from

Section 2 and assume that the sets G_t satisfy the following additional regularity properties (see the definition of coarse admissibility in [GN1]):

(CA1) for every bounded $\Omega \subset G$, there exists $c > 0$ such that

$$\Omega \cdot G_t \cdot \Omega \subset G_{t+c}$$

for all $t \geq t_0$.

(CA2) for every $c > 0$,

$$\sup_{t\geq t_0} \frac{\rho(H_{t+c})}{\rho(H_t)} < \infty.$$

Theorem 3.1. *Assume that properties* (CA1) *and* (CA2) *hold. Then*

(i) *For every $1 \leq p \leq \infty$, compact domain $D \subset X$ and $\phi \in L^p(D)$,*

$$\|\pi_X(\lambda_t)\phi\|_{L^p(D)} \ll_{p,D} \|\phi\|_{L^p(D)}.$$

(ii) *Let $1 \leq p \leq \infty$ and $l \in \mathbb{Z}_{\geq 0}$. Assume that for every compact domain $B \subset Y$, the averages β_t supprted on H_t satisfy the strong maximal inequality in $L^p_l(B)^+$ with respect to $\|\cdot\|_{L^p(B)}$, namely, for every $F \in L^p_l(B)^+$,*

$$\left\|\sup_{t\geq t_0} \pi_Y(\beta_t)F\right\|_{L^p(B)} \ll_{p,l,B} \|F\|_{L^p_l(B)}.$$

Then for every compact domain $D \subset X$, the family of measures λ_t satisfies the strong maximal inequality in $L^p_l(D)^+$ with respect to $\|\cdot\|_{L^p(D)}$, namely, for every $\phi \in L^p_l(D)^+$,

$$\left\|\sup_{t\geq t_0} \pi_X(\lambda_t)\phi\right\|_{L^p(D)} \ll_{p,l,D} \|\phi\|_{L^p_l(D)}.$$

We start by establishing a relation between L^p-norms on the spaces G and $Y = G/\Gamma$.

Lemma 3.2. *Let $1 \leq p \leq \infty$ and Ω be a compact domain of G. Then for every $F \in L^p(Y)$,*

$$\|F \circ \mathsf{p}_Y\|_{L^p(\Omega)} \ll_{p,\Omega} \|F\|_{L^p(\mathsf{p}_Y(\Omega))}.$$

Proof. We consider a measurable partition $\Omega = \sqcup_{i=1}^n \Omega_i$ such that the map p_Y is one-to-one on each Ω_i. Then by the triangle inequality and the definition of the measure on Y,

$$\|F \circ \mathsf{p}_Y\|_{L^p(\Omega)} \leq \sum_{i=1}^n \|(F \circ \mathsf{p}_Y)\chi_{\Omega_i}\|_{L^p(G)} = \sum_{i=1}^n \|F\chi_{\mathsf{p}_Y(\Omega_i)}\|_{L^p(Y)} \leq n\|F\|_{L^p(\mathsf{p}_Y(\Omega))}.$$

This proves the lemma. □

Given functions $\phi \in L^p_l(D)$ with $D \subset X$ and $\chi \in C^l_c(H)$ (the space of continuous functions with compact support and l continuous derivatives), we introduce functions $f : G \to \mathbb{R}$ and $F : Y \to \mathbb{R}$ defined by

$$f(g) := \chi(\mathsf{h}(g))\phi(\mathsf{p}_X(g)) \quad \text{and} \quad F(g\Gamma) := \sum_{\gamma\in\Gamma} f(g\gamma). \tag{3.1}$$

We note that if $\phi \in L^1(D)$, then it follows from (2.4) that $f \in L^1(G)$, and, hence, $F \in L^1(Y)$. The following lemma gives a more general estimate.

Lemma 3.3. *For every $1 \le p \le \infty$, $l \in \mathbb{N}_{\ge 0}$, and compact domain $D \subset X$,*

$$\|F\|_{L^p_l(Y)} \ll_{p,l,D,\chi} \|\phi\|_{L^p_l(D)}. \tag{3.2}$$

Proof. We first consider the case when $l = 0$. It follows from (2.2) that

$$\mathrm{supp}(f) \subset \Omega := \mathrm{supp}(\chi)^{-1}\mathsf{s}(D).$$

Recall that we choose the section s to be bounded on D, so that Ω is bounded. Since

$$f(h \cdot \mathsf{s}(x)) = \chi(h^{-1})\phi(x) \quad \text{for } (h, x) \in H \times X, \tag{3.3}$$

it follows from (2.4) that

$$\|f\|_{L^p(\Omega)} \ll_\chi \|\phi\|_{L^p(D)}. \tag{3.4}$$

Let $\Omega = \sqcup_{i=1}^n \Omega_i$ be a measurable partition such that the map p_Y is one-to-one on each Ω_i. We set $f_i = f\chi_{\Omega_i}$ and $F_i(g) = \sum_{\gamma\in\Gamma} f_i(g\gamma)$, where the sum is finite because f_i has compact support. Then

$$\|F_i\|_{L^p(Y)} = \|f_i\|_{L^p(\Omega)} \le \|f\|_{L^p(\Omega)}. \tag{3.5}$$

Since $F = \sum_{i=1}^n F_i$, estimate (3.2) now follows from (3.4) and (3.5).

Now let $l > 0$. In this case, G and H are assumed to be Lie groups, and every point in X has an open neighbourhood U with a smooth section $\mathsf{s} : U \to G$ of the factor map p_X. Using a partition of unity, we reduce the proof to the case when $\mathrm{supp}(\phi)$ is contained in one of these neighbourhoods U. Since the map

$$(\mathsf{h}, \mathsf{p}_X) : \mathsf{s}^{-1}(U) \to H \times U$$

defines a diffeomorphism on its domain, it follows from equation (3.3) that

$$\|f\|_{L^p_l(\Omega)} \ll_\chi \|\phi\|_{L^p_l(D)}. \tag{3.6}$$

Let $\{\psi_i\}_{j=1}^n$ be a partition of unity on G such that $\sum_{i=1}^n \psi_i = 1$ on Ω, and for every $i = 1, \dots n$, the map p_Y is a diffeomorphism on $\mathrm{supp}(\psi_i)$. Then for functions

$$F_i(g) := \sum_{\gamma\in\Gamma} (f\psi_i)(g\gamma),$$

we have

$$\|F_i\|_{L^p_l(Y)} = \|f\psi_i\|_{L^p_l(\Omega)} \ll \|f\|_{L^p_l(\Omega)} \|\psi_i\|_{C^l}. \tag{3.7}$$

Since $F = \sum_{i=1}^n F_i$, we deduce that

$$\|F\|_{L^p_l(Y)} \le \sum_{i=1}^n \|F_i\|_{L^p_l(Y)},$$

and the claim follows from (3.6) and (3.7). $\square$

3.2. **Coarse geometric comparison argument.**

Proof of Theorem 3.1. Writing $\phi = \phi^+ - \phi^-$ with $\phi^+ = \max(\phi, 0)$ and $\phi^- = \max(-\phi, 0)$, we reduce the proof of the first part of the theorem to the case when $\phi \in L^p(D)^+$.

Let $\phi \in L^p_l(D)^+$ and $\chi \in C^l_c(H)$ be a nonnegative function such that $\int_H \chi \, d\rho = 1$. We define functions $f : G \to \mathbb{R}$ and $F : Y \to \mathbb{R}$ as in (3.1). We note that by equation (3.3) we have

$$\operatorname{supp}(f) \subset \Omega := \operatorname{supp}(\chi)^{-1}\mathsf{s}(D) \quad \text{and} \quad \operatorname{supp}(F) \subset B := \mathsf{p}_Y(\Omega).$$

Since the section s is chosen to be bounded on D, it follows that Ω and B are bounded. By (2.4),

$$\|\pi_X(\lambda_t)\phi\|_{L^p(D)} \ll_D \|\pi_X(\lambda_t)(\phi) \circ \mathsf{p}_X\|_{L^p(\Omega)}, \tag{3.8}$$

and

$$\left\|\sup_{t \ge t_0} \pi_X(\lambda_t)\phi\right\|_{L^p(D)} \ll_D \left\|\sup_{t \ge t_0} \pi_X(\lambda_t)(\phi) \circ \mathsf{p}_X\right\|_{L^p(\Omega)} \tag{3.9}$$

We claim that there exists $c > 0$ (depending only on D and χ) such that for every $g \in \Omega$ and $t \ge t_0$,

$$\sum_{\gamma \in \Gamma_t} \phi(\mathsf{p}_X(g\gamma)) \le \int_{H_{t+c}} F(\mathsf{p}_Y(h^{-1}g))\, d\rho(h) = \int_{H_{t+c}} F(h^{-1}g\Gamma) d\rho(h). \tag{3.10}$$

Indeed, for $g \in \Omega$ and $\gamma \in \Gamma_t$ such that $\mathsf{p}_X(g\gamma) \in \operatorname{supp}(\phi) \subset D$, we have by (2.2),

$$\mathsf{h}(g\gamma)^{-1}\mathsf{s}(\mathsf{p}_X(g\gamma)) = g\gamma \in \Omega G_t.$$

Hence, by (CA1), there exists $c > 0$ such that

$$\mathsf{h}(g\gamma)^{-1} \operatorname{supp}(\chi) \subset \Omega G_t \mathsf{s}(D)^{-1} \operatorname{supp}(\chi) \subset G_{t+c},$$

and thus also

$$\operatorname{supp}(\chi) \subset \mathsf{h}(g\gamma) H_{t+c}.$$

Then we conclude using (2.3) and (3.1) that

$$\begin{aligned}
\phi(\mathsf{p}_X(g\gamma)) &= \phi(\mathsf{p}_X(g\gamma)) \int_{\mathsf{h}(g\gamma)H_{t+c}} \chi(h)\, d\rho(h) = \phi(\mathsf{p}_X(g\gamma)) \int_{H_{t+c}} \chi(\mathsf{h}(g\gamma)h)\, d\rho(h) \\
&= \int_{H_{t+c}} \chi(\mathsf{h}(h^{-1}g\gamma))\phi(\mathsf{p}_X(h^{-1}g\gamma))\, d\rho(h) = \int_{H_{t+c}} f(h^{-1}g\gamma)\, d\rho(h).
\end{aligned}$$

This implies that

$$\sum_{\gamma \in \Gamma_t} \phi(\mathsf{p}_X(g\gamma)) = \sum_{\gamma \in \Gamma_t} \int_{H_{t+c}} f(h^{-1}g\gamma)\, d\rho(h),$$

and since f is nonnegative, by 3.1 we have

$$\begin{aligned}
\sum_{\gamma \in \Gamma_t} \phi(\mathsf{p}_X(g\gamma)) &\le \sum_{\gamma \in \Gamma} \int_{H_{t+c}} f(h^{-1}g\gamma)\, d\rho(h) = \\
&= \int_{H_{t+c}} F(\mathsf{p}_Y(h^{-1}g))\, d\rho(h) = \int_{H_{t+c}} F(h^{-1}g\Gamma) d\rho(h).
\end{aligned}$$

This proves (3.10).

Now it follows from (3.10) and (CA2), using also (2.9) and (2.10) that on Ω,

$$\pi_X(\lambda_t)(\phi)\circ \mathsf{p}_X \le \left(\sup_{t\ge t_0}\frac{\rho(H_{t+c})}{\rho(H_t)}\right)\pi_Y(\beta_{t+c})F\circ \mathsf{p}_Y \ll \pi_Y(\beta_{t+c})F\circ \mathsf{p}_Y. \tag{3.11}$$

Hence, by (3.8), Lemma 3.2, Jensen's inequality, and Lemma 3.3,

$$\begin{aligned}\|\pi_X(\lambda_t)\phi\|_{L^p(D)} &\ll \|\pi_Y(\beta_{t+c})(F)\circ \mathsf{p}_Y\|_{L^p(\Omega)} \ll \|\pi_Y(\beta_{t+c})F\|_{L^p(B)}\\ &\le \|F\|_{L^p(B)} \ll \|\phi\|_{L^p(D)}.\end{aligned}$$

This proves the first part of the theorem.

To prove the second part of the theorem, we observe that by Lemma 3.3, we have $F\in L^p_l(B)^+$, and by the strong maximal inequality for the family β_t in $L^p_l(B)^+$ with respect to $\|\cdot\|_{L^p(B)}$, the function $\sup_{t\ge t_0}\pi_Y(\beta_{t+c})F$ is in $L^p(B)$. Hence, by (3.9), (3.11), and Lemma 3.2,

$$\left\|\sup_{t\ge t_0}\pi_X(\lambda_t)\phi\right\|_{L^p(D)} \ll \left\|\sup_{t\ge t_0}\pi_Y(\beta_{t+c})(F)\circ \mathsf{p}_Y\right\|_{L^p(\Omega)} \ll \left\|\sup_{t\ge t_0}\pi_Y(\beta_{t+c})F\right\|_{L^p(B)}.$$

Now the second part of the theorem follows from the strong maximal inequality in $L^p_l(B)^+$ combined with Lemma 3.3. □

4. The mean ergodic theorem

4.1. **Average admissibility of the restricted sets.** The goal of this section is prove the mean ergodic theorem for the averages $\pi_X(\lambda_t)$ defined in (2.9), based on the mean ergodic theorems for the averages $\pi_Y(\beta_t)$ defined in (2.10). We shall use notation from Section 2 and assume that the sets G_t satisfy additional regularity properties, as follows. Note that properties (A2) and (A2') depend on a parameter $r\in(1,\infty)$, but for simplicity we suppress this dependence in the notation.

(A1) For every $\varepsilon\in(0,1)$, there exists a neighbourhood O_ε of identity in G such that

$$O_\varepsilon\cdot G_t\cdot O_\varepsilon\subset G_{t+\varepsilon}$$

for all $t\ge t_0$.

(A2) For every $u\in G$, and for every compact $\Omega\subset G$, there exists $\omega(\varepsilon)>0$ that converges to 0 as $\varepsilon\to 0^+$ such that

$$\left(\int_\Omega (\rho(H_{t+\varepsilon}[u,v])-\rho(H_t[u,v]))^r\,dm(v)\right)^{1/r}\le \omega(\varepsilon)\rho(H_t)$$

for all $t\ge t_0$ and $\varepsilon\in(0,1)$.

(A2′) For every compact $\Omega\subset G$, there exists $\omega(\varepsilon)>0$ that converges to 0 as $\varepsilon\to 0^+$ such that

$$\left(\int_{\Omega\times\Omega} (\rho(H_{t+\varepsilon}[u,v])-\rho(H_t[u,v]))^r\,dm(u)dm(v)\right)^{1/r}\le \omega(\varepsilon)\rho(H_t).$$

for all $t\ge t_0$ and $\varepsilon\in(0,1)$.

(A3) For every $g_1, g_2 \in G$, the limit

$$\Theta(g_1, g_2) := \lim_{t\to\infty} \frac{\rho(H_t[g_1, g_2])}{\rho(H_t)}$$

exists.

(S) every $x \in X$ has a neighborhood U such that there exists a continuous section $\mathsf{s} : U \to G$ of the factor map p_X.

Property (A2′) is used only in the proof of the mean ergodic theorem (Theorem 4.1), and property (A2) is used only in the proof of the pointwise ergodic theorem (Theorem 5.1).

Theorem 4.1. *Let $1 \le p < \infty$. Assume that* (CA1), (CA2), (A1), (A2′) *with some $r \ge p$,* (A3), (S) *hold, and for every $g_1, g_2 \in G$ and every compact $B \subset Y$, the family $\beta_t^{g_1,g_2}$ satisfies the mean ergodic theorem in $L^p(B)$, namely, for every $F \in L^p(B)$,*

$$\left\| \pi_Y(\beta_t^{g_1,g_2})F - \int_Y F\, d\mu \right\|_{L^p(B)} \to 0 \quad \text{as } t \to \infty.$$

Then for every compact domain $D \subset X$, the family λ_t satisfies the mean ergodic theorem in $L^p(D)$, namely, for every $\phi \in L^p(D)$ and for almost every $x \in D$

$$\left\| \pi_X(\lambda_t)\phi(x) - \int_X \phi\, d\nu_x \right\|_{L^p(D)} \to 0 \quad \text{as } t \to \infty, \tag{4.1}$$

where ν_x is the measure defined in (2.7).

Besides the discrete averages $\pi_X(\lambda_t)$ supported on Γ, it will be also convenient to consider their continuous analogues which are defined by

$$\pi_X(\lambda_t^G)\phi(x) := \frac{1}{\rho(H_t)} \int_{G_t} \phi(xg)dm(g), \quad \phi \in L^p(D), \tag{4.2}$$

for every $x \in X$. It would be convenient in the proof to use averages without normalisation. Namely, we set

$$\pi_X(\Lambda_t)\phi(x) := \sum_{\gamma\in\Gamma\cap G_t} \phi(x\gamma) \quad \text{and} \quad \pi_X(\Lambda_t^G)\phi(x) := \int_{G_t} \phi(xg)dm(g). \tag{4.3}$$

We start the proof of Theroem 4.1 with a lemma:

Lemma 4.2. *Let D be a compact domain in X, $1 \le p < q \le \infty$, and $\phi \in L^q(D)$.*

(i) *Suppose that* (A2) *holds with $r = q/(q-1)$ (here $r = 1$ if $q = \infty$). Then for every $x \in X$, the estimate*

$$\left|\pi_X(\Lambda_{t+\varepsilon}^G)\phi(x) - \pi_X(\Lambda_t^G)\phi(x)\right| \ll_{q,x,D} \omega(\varepsilon)\rho(H_t)\|\phi\|_{L^q(D)}$$

holds for all $t \ge t_0$ and $\varepsilon \in (0, 1)$.

(ii) *Suppose that* (A2′) *holds where $r = pq/(q-p)$ (here $r = p$ if $q = \infty$). Then the estimate*

$$\left\|\pi_X(\Lambda_{t+\varepsilon}^G)\phi - \pi_X(\Lambda_t^G)\phi\right\|_{L^p(D)} \ll_{p,q,D} \omega(\varepsilon)\rho(H_t)\|\phi\|_{L^q(D)}$$

holds for all $t \ge t_0$ and $\varepsilon \in (0, 1)$.

Proof. To prove (i), we are required to estimate

$$\left|\int_{G_{t+\varepsilon}} \phi(xg)\, dm(g) - \int_{G_t} \phi(xg)\, dm(g)\right| = \left|\int_{G_{t+\varepsilon}-G_t} \phi(xg)\, dm(g)\right|.$$

It follows from invariance of m, the equivariance of p_X and (2.4) that

$$\begin{aligned}\int_{G_{t+\varepsilon}-G_t} \phi(xg)\, dm(g) &= \int_{G_{t+\varepsilon}-G_t} \phi(\mathsf{p}_X(\mathsf{s}(x)g))\, dm(g) \\ &= \int_{\mathsf{s}(x)(G_{t+\varepsilon}-G_t)} \phi(\mathsf{p}_X(g))\, dm(g) \\ &= \int_{(y,h):\, \mathsf{s}(x)^{-1}h\mathsf{s}(y)\in G_{t+\varepsilon}-G_t} \phi(\mathsf{p}_X(h\mathsf{s}(y)))\, d\rho(h) d\xi(y) \\ &= \int_X \phi(y)\left(\rho(H_{t+\varepsilon}[\mathsf{s}(x),\mathsf{s}(y)]) - \rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])\right) d\xi(y),\end{aligned} \tag{4.4}$$

Hence, by Hölder's inequality,

$$\begin{aligned}&\left|\int_{G_{t+\varepsilon}-G_t} \phi(xg)\, dm(g)\right| \\ \le&\|\phi\|_{L^q(D)}\left(\int_D \left(\rho(H_{t+\varepsilon}[\mathsf{s}(x),\mathsf{s}(y)]) - \rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])\right)^r\, d\xi(y)\right)^{1/r},\end{aligned}$$

where $r = q/(q-1)$ is the exponent conjugate to q. We pick a compact set O of H with positive measure and put $\Omega = O\mathsf{s}(D)$. Then it follows from (2.4) and (A2) that since $\mathsf{s}(x)$ and $\mathsf{s}(y)$ vary in a compact set

$$\begin{aligned}&\int_D \left(\rho(H_{t+\varepsilon}[\mathsf{s}(x),\mathsf{s}(y)]) - \rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])\right)^r\, d\xi(y) \\ \ll_O\ & \int_\Omega \left(\rho(H_{t+\varepsilon}[\mathsf{s}(x),v]) - \rho(H_t[\mathsf{s}(x),v])\right)^r\, dm(v) \\ \ll\ & \omega(\varepsilon)^r\rho(H_t)^r.\end{aligned}$$

This implies the first part of the lemma.

To prove the second part, we are required to estimate

$$\begin{aligned}&\int_D \left|\int_{G_{t+\varepsilon}-G_t} \phi(xg)\, dm(g)\right|^p\, d\xi(x) \\ =&\int_D \left|\int_X \phi(y)\left(\rho(H_{t+\varepsilon}[\mathsf{s}(x),\mathsf{s}(y)]) - \rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])\right) d\xi(y)\right|^p\, d\xi(x)\end{aligned}$$

Since $\mathrm{supp}(\phi) \subset D$, it follows from (4.4) and Hölder's inequality that

$$\begin{aligned}&\left|\int_{G_{t+\varepsilon}-G_t} \phi(xg)\, dm(g)\right|^p \\ \le&\xi(D)^{p(1-1/p)}\int_D |\phi(y)|^p\left(\rho(H_{t+\varepsilon}[\mathsf{s}(x),\mathsf{s}(y)]) - \rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])\right)^p\, d\xi(y).\end{aligned}$$

Moreover, integrating over $x \in D$ w.r.t. ξ and applying Hölder's inequality one more time, we deduce that

$$\int_{D\times D} |\phi(y)|^p \left(\rho(H_{t+\varepsilon}[\mathsf{s}(x),\mathsf{s}(y)]) - \rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])\right)^p \, d\xi(x)d\xi(y)$$

$$\leq \xi(D)^{p/q}\|\phi\|^p_{L^q(D)} \left(\int_{D\times D} \left(\rho(H_{t+\varepsilon}[\mathsf{s}(x),\mathsf{s}(y)]) - \rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])\right)^{ps} \, d\xi(x)d\xi(y)\right)^{1/s},$$

where $s = q/(q-p)$ is the exponent conjugate to q/p (here $s = 1$ if $q = \infty$). Hence, we conclude that

$$\int_D \left|\int_{G_{t+\varepsilon}-G_t} \phi(xg)\, dm(g)\right|^p \, d\xi(x)$$

$$\ll \|\phi\|^p_{L^q(D)} \left(\int_{D\times D} \left(\rho(H_{t+\varepsilon}[\mathsf{s}(x),\mathsf{s}(y)]) - \rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])\right)^{ps} \, d\xi(x)d\xi(y)\right)^{1/s}.$$

As in the first part of the argument, we pick a compact set O of H with positive measure and set $\Omega = O\mathsf{s}(D)$. Then it follows from (2.4) and (A2$'$) that

$$\begin{aligned}
&\int_{D\times D} \left(\rho(H_{t+\varepsilon}[\mathsf{s}(x),\mathsf{s}(y)]) - \rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])\right)^{ps} \, d\xi(x)d\xi(y)\\
\ll_O &\int_{\Omega\times\Omega} \left(\rho(H_{t+\varepsilon}[u,v]) - \rho(H_t[u,v])\right)^{ps} \, dm(u)dm(v)\\
\ll\ &\omega(\varepsilon)^{ps}\rho(H_t)^{ps}.
\end{aligned}$$

This implies the required estimate. □

The proof of Theorem 4.1 now continues with

4.2. **Geometric comparison argument.** We first observe that for every $\phi, \psi \in L^p(D)$ and $x \in X$ we have

$$\|\pi_X(\lambda_t)\phi - \pi_X(\lambda_t)\psi\|_{L^p(D)} \ll \|\phi-\psi\|_{L^p(D)},$$

$$\left\|\int_X \phi\, d\nu_x - \int_X \psi\, d\nu_x\right\|_{L^p(D)} \ll \|\phi-\psi\|_{L^p(D)}.$$

The first estimate is proved in Theorem 3.1(i). To prove the second estimate we observe that it follows from (CA1)–(CA2) that the density $\Theta(x,\cdot)$ of the measure ν_x is uniformly bounded on D, with the bound uniform as x varies in compact sets in G. Hence, since $\xi(D) < \infty$, the second estimate follows from Hölder's inequality, by definition of ν_x.

The above estimates imply that it is sufficient to verify (4.1) for a dense family of functions in $L^p(D)$, and we shall prove that (4.1) holds for $\phi \in L^q(D)$ with $q > p$ such that $r = pq/(q-p)$. Since every such ϕ can be written as $\phi = \phi^+ - \phi^-$ with $\phi^+ \geq 0$ and $\phi^- \geq 0$ are in $L^q(D)$, the proof reduces to the case when $\phi \geq 0$. Moreover, because condition (S) is satisfied, decomposing ϕ as a finite sum of functions with small supports, we reduce the proof to the situation when there exists a section $\mathsf{s} : X \to G$ of the factor map $\mathsf{p}_X : G \to H\backslash G = X$ such that $\mathsf{s}|_D$ is continuous.

Using (2.4), we deduce that for every $x \in X$, we have

$$\begin{aligned}\pi_X(\lambda_t^G)\phi(x) &= \frac{1}{\rho(H_t)}\int_{G_t} \phi(\mathsf{p}_X(\mathsf{s}(x)g))\, dm(g)\\ &= \frac{1}{\rho(H_t)}\int_{(y,h):\, \mathsf{s}(x)^{-1}h\mathsf{s}(y)\in G_t} \phi(\mathsf{p}_X(h\mathsf{s}(y)))\, d\rho(h)d\xi(y)\\ &= \int_D \phi(y)\frac{\rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])}{\rho(H_t)}\, d\xi(y).\end{aligned}$$

By (CA1)–(CA2), $\frac{\rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])}{\rho(H_t)}$ is bounded uniformly when x is fixed and y varies on D. Hence, it follows from (A3) and the dominated convergence theorem (since ϕ is certainly in $L^1(D)$) , that for every $x \in X$,

$$\pi_X(\lambda_t^G)\phi(x) \to \int_D \phi\, d\nu_x = \int_X \phi\, d\nu_x \quad \text{as } t \to \infty. \tag{4.5}$$

To obtain convergence in $L^p(D)$, consider the difference

$$\left|\pi_X(\lambda_t^G)\phi(x) - \int_D \phi\, d\nu_x\right|^p = \left|\int_D \phi(y)\left(\frac{\rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])}{\rho(H_t)} - \Theta(\mathsf{s}(x),\mathsf{s}(y))\right) d\xi(y)\right|^p$$

Therefore, applying the dominated convergence theorem one more time, and the fact that $\phi \in L^p(D)$, we deduce by integrating over $x \in D$ that the convergence also holds in $L^p(D)$.

To conclude the proof of Theorem 4.1, it remains to show that

$$\left\|\pi_X(\lambda_t)\phi - \pi_X(\lambda_t^G)\phi\right\|_{L^p(D)} \to 0 \quad \text{as } t \to \infty \tag{4.6}$$

for every ϕ in the dense subset of $L^p(D)^+$ we chose, namely every $\phi \in L^q(D)^+$. This calls for a comparison between the discrete average supported on $x \cdot \Gamma_t$, and the continuous averages supported on $x \cdot G_t$. We will therefore apply local analysis to effect this comparison.

Let $\varepsilon \in (0, 1/42)$ and O be a compact symmetric neighborhood of identity in G such that

$$O \cdot G_t \cdot g^{-1}Og \subset G_{t+\varepsilon} \quad \text{for every } t \geq t_0 \text{ and } g \in \mathsf{s}(D), \tag{4.7}$$

$$G_t \cdot \mathsf{s}(yO)^{-1} \subset G_{t+\varepsilon} \cdot \mathsf{s}(y)^{-1} \quad \text{for every } t \geq t_0 \text{ and } y \in D. \tag{4.8}$$

Such a neighbourhood exists by (A1) and continuity of s on D. Let $\chi \in C_c(H)^+$ with $\mathrm{supp}(\chi) \subset O \cap H$ be normalized so that $\int_H \chi\, d\rho = 1$, and let $f : G \to \mathbb{R}$ be defined as in (3.1).

We claim that for $u \in G$,

$$\pi_X(\Lambda_t)\phi(\mathsf{p}_X(u)) \leq \sum_{\gamma\in\Gamma}\int_{H_{t+\varepsilon}[u,\mathsf{r}(u\gamma)]} f(h^{-1}u\gamma)\, d\rho(h), \tag{4.9}$$

and

$$\pi_X(\Lambda_t)\phi(\mathsf{p}_X(u)) \geq \sum_{\gamma\in\Gamma}\int_{H_{t-\varepsilon}[u,\mathsf{r}(u\gamma)]} f(h^{-1}u\gamma)\, d\rho(h). \tag{4.10}$$

To establish (4.9), it suffices to consider $\gamma \in \Gamma_t$ satisfying $\mathsf{p}_X(u\gamma) \in \mathrm{supp}(\phi) \subset D$, where we have by (2.2),

$$\mathsf{h}(u\gamma)^{-1}\mathsf{r}(u\gamma) = u\gamma \in uG_t,$$

and by (4.7),

$$\mathsf{h}(u\gamma)^{-1}\mathsf{r}(u\gamma)\left(\mathsf{r}(u\gamma)^{-1}O\mathsf{r}(u\gamma)\right) = \mathsf{h}(u\gamma)^{-1}O\mathsf{r}(u\gamma) \subset uG_t \cdot \mathsf{r}(u\gamma)^{-1}O\mathsf{r}(u\gamma) \subset uG_{t+\varepsilon}.$$

Hence,

$$\mathsf{h}(u\gamma)^{-1}O \subset uG_{t+\varepsilon}\mathsf{r}(u\gamma)^{-1}$$

and thus also

$$\mathrm{supp}(\chi) \subset \mathsf{h}(u\gamma)H_{t+\varepsilon}[u, \mathsf{r}(u\gamma)].$$

Therefore by invariance of ρ and (2.3),

$$\begin{aligned}
\phi(\mathsf{p}_X(u\gamma)) &= \phi(\mathsf{p}_X(u\gamma)) \int_{\mathsf{h}(u\gamma)H_{t+\varepsilon}[u,\mathsf{r}(u\gamma)]} \chi(h)\, d\rho(h) \\
&= \phi(\mathsf{p}_X(u\gamma)) \int_{H_{t+\varepsilon}[u,\mathsf{r}(u\gamma)]} \chi(\mathsf{h}(u\gamma)h)\, d\rho(h) \\
&= \int_{H_{t+\varepsilon}[u,\mathsf{r}(u\gamma)]} \chi(\mathsf{h}(h^{-1}u\gamma))\phi(\mathsf{p}_X(h^{-1}u\gamma))\, d\rho(h) \\
&= \int_{H_{t+\varepsilon}[u,\mathsf{r}(u\gamma)]} f(h^{-1}u\gamma)\, d\rho(h),
\end{aligned}$$

and we conclude that

$$\pi_X(\Lambda_t)\phi(\mathsf{p}_X(u)) \le \sum_{\gamma\in\Gamma_t} \int_{H_{t+\varepsilon}[u,\mathsf{r}(u\gamma)]} f(h^{-1}u\gamma)\, d\rho(h),$$

Since $f \ge 0$, we can sum over all $\gamma \in \Gamma$, and this implies (4.9).

To prove (4.10), we observe that for $\gamma \in \Gamma - \Gamma_t$ such that $\mathsf{p}_X(u\gamma) \in \mathrm{supp}(\phi) \subset D$, we have

$$\mathrm{supp}(\chi) \cap \mathsf{h}(u\gamma)H_{t-\varepsilon}[u, \mathsf{r}(u\gamma)] = \emptyset. \tag{4.11}$$

Indeed, if $h \in \mathrm{supp}(\chi) \subset O$ belongs to this intersection, then

$$u^{-1}\mathsf{h}(u\gamma)^{-1}h\mathsf{r}(u\gamma) \in G_{t-\varepsilon},$$

and by (2.2) and (4.7),

$$\gamma = u^{-1}\mathsf{h}(u\gamma)^{-1}\mathsf{r}(u\gamma) \in G_{t-\varepsilon} \cdot \mathsf{r}(u\gamma)^{-1}h^{-1}\mathsf{r}(u\gamma) \subset G_t,$$

which gives a contradiction. Thus we can now deduce from (4.11) that

$$\begin{aligned}
&\sum_{\gamma\in\Gamma}\int_{H_{t-\varepsilon}[u,\mathsf{r}(u\gamma)]} f(h^{-1}u\gamma)\,d\rho(h)\\
=&\sum_{\gamma\in\Gamma}\int_{H_{t-\varepsilon}[u,\mathsf{r}(u\gamma)]} \chi(\mathsf{h}(h^{-1}u\gamma))\phi(\mathsf{p}_X(hu\gamma))\,d\rho(h)\\
=&\sum_{\gamma\in\Gamma}\phi(\mathsf{p}_X(u\gamma))\int_{\mathsf{h}(u\gamma)H_{t-\varepsilon}[u,\mathsf{r}(u\gamma)]}\chi(h)\,d\rho(h)\\
=&\sum_{\gamma\in\Gamma_t}\phi(\mathsf{p}_X(u\gamma))\int_{\mathsf{h}(u\gamma)H_{t-\varepsilon}[u,\mathsf{r}(u\gamma)]}\chi\,d\rho\le\sum_{\gamma\in\Gamma_t}\phi(\mathsf{p}_X(u\gamma))=\pi_X(\Lambda_t)\phi(\mathsf{p}_X(u)),
\end{aligned}$$

which proves (4.10).

A similar line of reasoning also yields the analogous estimates for the averages $\pi_X(\Lambda_t^G)$. Namely, for $u\in G$

$$\pi_X(\Lambda_t^G)\phi(\mathsf{p}_X(u))\le\int_G\int_{H_{t+\varepsilon}[u,\mathsf{r}(ug)]} f(h^{-1}ug)\,d\rho(h)dm(g),\tag{4.12}$$

and

$$\pi_X(\Lambda_t^G)\phi(\mathsf{p}_X(u))\ge\int_G\int_{H_{t-\varepsilon}[u,\mathsf{r}(ug)]} f(h^{-1}ug)\,d\rho(h)dm(g).\tag{4.13}$$

4.3. **Local analysis.** Let us now begin with the next step of the argument. Let $\Omega:=\mathrm{supp}(\chi)^{-1}\mathsf{s}(D)$ and $B=\mathsf{p}_Y(\Omega)$. Then $\mathrm{supp}(f)\subset\Omega$, and there exists a finite subset $\{v_i\}_{i=1}^n\subset\Omega$ such that

$$\Omega\subset\bigcup_{i=1}^n v_iO.$$

Then it follows from (4.8) that for every $g\in Hv_iO$,

$$H_{t-\varepsilon}[u,\mathsf{r}(v_i)]\subseteq H_t[u,\mathsf{r}(g)]\subseteq H_{t+\varepsilon}[u,\mathsf{r}(v_i)].\tag{4.14}$$

We fix a finite measurable partition

$$\mathrm{supp}(f)=\bigsqcup_{i=1}^n C_i\quad\text{such that } C_i\subset v_iO.$$

Let

$$f_i=f\cdot\chi_{C_i}\quad\text{and}\quad F_i(\mathsf{p}_Y(g))=\sum_{\gamma\in\Gamma}f_i(g\gamma),$$

where the sum is finite because f_i has compact support. We note that by the definition of μ, we have $\int_Y F_i\,d\mu=\int_X f_i\,dm$, and by Lemma 3.3, $F_i\in L^q(B)$. Now

we deduce from (4.14) that

$$\begin{aligned}\sum_{\gamma\in\Gamma}\int_{H_t[u,\mathsf{r}(u\gamma)]} f(h^{-1}u\gamma)\,d\rho(h) &= \sum_{i=1}^n\sum_{\gamma\in\Gamma}\int_{H_t[u,\mathsf{r}(u\gamma)]} f_i(h^{-1}u\gamma)\,d\rho(h)\\ &\le \sum_{i=1}^n\sum_{\gamma\in\Gamma}\int_{H_{t+\varepsilon}[u,\mathsf{r}(v_i)]} f_i(h^{-1}u\gamma)\,d\rho(h)\\ &= \sum_{i=1}^n\int_{H_{t+\varepsilon}[u,\mathsf{r}(v_i)]} F_i(h^{-1}\mathsf{p}_Y(u))\,d\rho(h). \end{aligned}\tag{4.15}$$

A similar argument also gives the lower estimate

$$\sum_{\gamma\in\Gamma}\int_{H_t[u,\mathsf{r}(u\gamma)]} f(h^{-1}u\gamma)\,d\rho(h) \ge \sum_{i=1}^n\int_{H_{t-\varepsilon}[u,\mathsf{r}(v_i)]} F_i(h^{-1}\mathsf{p}_Y(u))\,d\rho(h). \tag{4.16}$$

There exists a finite subset $\{u_j\}_{j=1}^m\subset\Omega$ and a finite measurable partition

$$\Omega=\bigsqcup_{j=1}^m\Omega_j\quad\text{such that } \Omega_j\subset u_jO.$$

By (4.7), for every $u\in u_jO$,

$$H_{t-\varepsilon}[u_j,\mathsf{r}(v_i)]\subseteq H_t[u,\mathsf{r}(v_i)]\subseteq H_{t+\varepsilon}[u_j,\mathsf{r}(v_i)].$$

We introduce a measurable function $\mathcal{A}_t:\Omega\to\mathbb{R}$ which is defined piecewise by

$$\mathcal{A}_t(u):=\sum_{i=1}^n\int_{H_t[u_j,\mathsf{r}(v_i)]} F_i(h^{-1}\mathsf{p}_Y(u))\,d\rho(h)\quad\text{for } u\in\Omega_j. \tag{4.17}$$

Then for $u\in\Omega$,

$$\mathcal{A}_{t-\varepsilon}(u)\le\sum_{i=1}^n\int_{H_t[u,\mathsf{r}(v_i)]} F_i(h^{-1}\mathsf{p}_Y(u))\,d\rho(h)\le\mathcal{A}_{t+\varepsilon}(u),$$

and combining (4.9), (4.10), (4.15), (4.16), we deduce that for $u\in\Omega$,

$$\mathcal{A}_{t-3\varepsilon}(u)\le\pi_X(\Lambda_t)\phi(\mathsf{p}_X(u))\le\mathcal{A}_{t+3\varepsilon}(u). \tag{4.18}$$

4.4. **The duality argument.** We would now like to exploit the information regarding the ergodic behavior of the H-orbits in G/Γ, and our next task is to prove a similar estimate for the averages $\pi_X(\Lambda_t^G)$. For $u\in\Omega_j$, we have by (4.14) and invariance of m,

$$\begin{aligned}\int_G\int_{H_t[u,\mathsf{r}(ug)]} f(h^{-1}ug)\,d\rho(h)dm(g) &\le \sum_{i=1}^n\int_G\int_{H_{t+\varepsilon}[u,\mathsf{r}(v_i)]} f_i(h^{-1}ug)\,d\rho(h)dm(g)\\ &= \sum_{i=1}^n\rho(H_{t+\varepsilon}[u,\mathsf{r}(v_i)])\int_G f_i\,dm\\ &\le \sum_{i=1}^n\rho(H_{t+2\varepsilon}[u_j,\mathsf{r}(v_i)])\int_G f_i\,dm,\end{aligned}$$

and similarly,

$$\int_G \int_{H_t[u,\mathsf{r}(ug)]} f(h^{-1}ug)\, d\rho(h) dm(g) \geq \sum_{i=1}^n \rho(H_{t-2\varepsilon}[u_j, \mathsf{r}(v_i)]) \int_G f_i \, dm.$$

As in (4.17), we introduce a function $\mathcal{A}_t^G : \Omega \to \mathbb{R}$ defined by

$$\mathcal{A}_t^G(u) = \sum_{i=1}^n \rho(H_t[u_j, \mathsf{r}(v_i)]) \int_G f_i \, dm \quad \text{for } u \in \Omega_j. \tag{4.19}$$

Then it follows from the above estimates combined with (4.12) and (4.13) that for $u \in \Omega$,

$$\mathcal{A}_{t-3\varepsilon}^G(u) \leq \pi_X(\Lambda_t^G)\phi(\mathsf{p}_X(u)) \leq \mathcal{A}_{t+3\varepsilon}^G(u). \tag{4.20}$$

By (2.4) and Lemma 4.2(ii) combined with (A2′), we have

$$\left\|\pi_X(\Lambda_{t+\varepsilon}^G)\phi \circ \mathsf{p}_X - \pi_X(\Lambda_t^G)\phi \circ \mathsf{p}_X\right\|_{L^p(\Omega)} \leq \omega(\varepsilon)\rho(H_t)\|\phi\|_{L^q(D)}, \tag{4.21}$$

where $\omega(\varepsilon) \to 0$ as $\varepsilon \to 0^+$. Combining (4.20) and (4.21), we deduce that

$$\begin{aligned} \left\|\mathcal{A}_{t+\varepsilon}^G - \mathcal{A}_t^G\right\|_{L^p(\Omega)} &\leq \left\|\pi_X(\Lambda_{t+4\varepsilon}^G)\phi \circ \mathsf{p}_X - \pi_X(\Lambda_{t-3\varepsilon}^G)\phi \circ \mathsf{p}_X\right\|_{L^p(\Omega)} \\ &\ll \omega(7\varepsilon)\rho(H_t)\|\phi\|_{L^q(D)}. \end{aligned} \tag{4.22}$$

Now we show that

$$\|\mathcal{A}_t - \mathcal{A}_t^G\|_{L^p(\Omega)} = o(\rho(H_t)) \quad \text{as } t \to \infty. \tag{4.23}$$

This is where we utilise the mean ergodic theorem for H acting in $L^p(B) \subset L^p(G/\Gamma)$. Using the triangle inequality, Lemma 3.2, and (CA1)–(CA2), we obtain

$$\begin{aligned} &\|\mathcal{A}_t - \mathcal{A}_t^G\|_{L^p(\Omega)} \leq \sum_{j=1}^m \|\mathcal{A}_t - \mathcal{A}_t^G\|_{L^p(\Omega_j)} \\ &= \sum_{j=1}^m \left\| \sum_{i=1}^n \int_{H_t[u_j,\mathsf{r}(v_i)]} F_i(h^{-1}\mathsf{p}_Y(g))\, d\rho(h) - \sum_{i=1}^n \rho(H_t[u_j, \mathsf{r}(v_i)]) \int_G f_i \, dm \right\|_{L^p(\Omega_j)} \\ &\ll \sum_{i=1}^n \sum_{j=1}^m \rho(H_t[u_j, \mathsf{r}(v_i)]) \left\|\mathcal{E}_t^{ij}\right\|_{L^p(B)} \ll \rho(H_t) \sum_{i=1}^n \sum_{j=1}^m \left\|\mathcal{E}_t^{ij}\right\|_{L^p(B)}, \end{aligned}$$

where

$$\mathcal{E}_t^{ij}(y) := \left| \frac{1}{\rho(H_t[u_j, \mathsf{r}(v_i)])} \int_{H_t[u_j,\mathsf{r}(v_i)]} F_i(h^{-1}y)\, d\rho(h) - \int_Y F_i \, d\mu \right|.$$

Since $\mathcal{E}_t^{ij} \to 0$ in $L^p(B)$ as $t \to \infty$ by the mean ergodic theorem for the family $\beta_t^{g_1,g_2}$, this proves (4.23).

Using that

$$\begin{aligned} \|\mathcal{A}_{t+\varepsilon} - \mathcal{A}_t\|_{L^p(\Omega)} \leq &\|\mathcal{A}_{t+\varepsilon} - \mathcal{A}_{t+\varepsilon}^G\|_{L^p(\Omega)} + \|\mathcal{A}_{t+\varepsilon}^G - \mathcal{A}_t^G\|_{L^p(\Omega)} \\ &+ \|\mathcal{A}_t^G - \mathcal{A}_t\|_{L^p(\Omega)}. \end{aligned}$$

we deduce from (4.22) and (4.23) that

$$\limsup_{t\to\infty} \frac{\|\mathcal{A}_{t+\varepsilon} - \mathcal{A}_t\|_{L^p(\Omega)}}{\rho(H_t)} \ll \omega(7\varepsilon)\|\phi\|_{L^q(D)}. \tag{4.24}$$

Now we finally complete the proof by showing that the difference between the discrete sampling operators and the continuous ones converges to zero in norm, namely :

$$\left\|\pi_X(\Lambda_t)\phi - \pi_X(\Lambda_t^G)\phi\right\|_{L^p(D)} = o(\rho(H_t)) \quad \text{as } t\to\infty.$$

By (2.4),

$$\left\|\pi_X(\Lambda_t)\phi - \pi_X(\Lambda_t^G)\phi\right\|_{L^p(D)} \ll \left\|\pi_X(\Lambda_t)\phi\circ \mathsf{p}_X - \pi_X(\Lambda_t^G)\phi\circ \mathsf{p}_X\right\|_{L^p(\Omega)}.$$

By (4.18),

$$\begin{aligned}&\left\|\pi_X(\Lambda_t)\phi\circ \mathsf{p}_X - \pi_X(\Lambda_t^G)\phi\circ \mathsf{p}_X\right\|_{L^p(\Omega)}\\ &\le \left\|\pi_X(\Lambda_t)\phi\circ \mathsf{p}_X - \mathcal{A}_{t-3\varepsilon}\right\|_{L^p(\Omega)} + \left\|\mathcal{A}_{t-3\varepsilon} - \pi_X(\Lambda_t^G)\phi\circ \mathsf{p}_X\right\|_{L^p(\Omega)}\\ &\le \left\|\mathcal{A}_{t+3\varepsilon} - \mathcal{A}_{t-3\varepsilon}\right\|_{L^p(\Omega)} + \left\|\mathcal{A}_{t-3\varepsilon} - \pi_X(\Lambda_t^G)\phi\circ \mathsf{p}_X\right\|_{L^p(\Omega)},\end{aligned}$$

and by (4.20),

$$\begin{aligned}&\left\|\mathcal{A}_{t-3\varepsilon} - \pi_X(\Lambda_t^G)\phi\circ \mathsf{p}_X\right\|_{L^p(\Omega)}\\ &\le \left\|\mathcal{A}_{t-3\varepsilon} - \mathcal{A}^G_{t-3\varepsilon}\right\|_{L^p(\Omega)} + \left\|\pi_X(\Lambda_t^G)\phi\circ \mathsf{p}_X - \mathcal{A}^G_{t-3\varepsilon}\right\|_{L^p(\Omega)}\\ &\le \left\|\mathcal{A}_{t-3\varepsilon} - \mathcal{A}^G_{t-3\varepsilon}\right\|_{L^p(\Omega)} + \left\|\pi_X(\Lambda_t^G)\phi\circ \mathsf{p}_X - \pi_X(\Lambda^G_{t-6\varepsilon})\phi\circ \mathsf{p}_X\right\|_{L^p(\Omega)}.\end{aligned}$$

Therefore, it follows from (4.24), (4.23), and (4.21) that

$$\limsup_{t\to\infty}\frac{\left\|\pi_X(\Lambda_t)\phi\circ \mathsf{p}_X - \pi_X(\Lambda_t^G)\phi\circ \mathsf{p}_X\right\|_{L^p(\Omega)}}{\rho(H_t)} \ll (\omega(42\varepsilon)+\omega(6\varepsilon))\|\phi\|_{L^q(D)}.$$

Since $\omega(\varepsilon)\to 0$ as $\varepsilon\to 0^+$, this completes the proof of Theorem 4.1.

5. The pointwise ergodic theorem

We now prove that the validity of the the pointwise ergodic theorem for $\pi_Y(\beta_t)$ implies its validity for $\pi_X(\lambda_t)$.

Theorem 5.1. *Let $1 < p \le \infty$. Assume that* (CA1), (CA2), (A1), (A2) *with $r = p/(p-1)$,* (A3), and (S) *hold. Assume that for every $g_1, g_2 \in G$ and every compact domain $B \subset Y$, family $\beta_t^{g_1,g_2}$ satisfies the pointwise ergodic theorem in $L^p(B)$, namely, for every $F \in L^p(B)$,*

$$\lim \pi_Y(\beta_t^{g_1,g_2})F(y) \to \int_Y F\,d\mu \quad \textit{for almost every } y\in B.$$

Then for every compact domain $D \subset X$, the family λ_t satisfies the pointwise ergodic theorem in $L^p(D)$, namely, for every $\phi \in L^p(D)$,

$$\lim_{t\to\infty}\pi_X(\lambda_t)\phi(x) = \int_X \phi\,d\nu_x \quad \textit{for almost every } x\in D. \tag{5.1}$$

Proof. In the proof we shall use notations introduced in the proof of Theorem 4.1. As in that proof, we reduce our argument to the case when $\phi \ge 0$ and the section s is continuous on D. Moreover, because of (4.5), it is sufficient to show that

$$\lim_{t\to\infty}(\pi_X(\lambda_t)\phi(x) - \pi_X(\lambda_t^G)\phi(x)) = 0 \quad \text{for almost every } x\in D.$$

Let $\varepsilon \in (0, 1/12)$, $\Omega = D_0\mathsf{s}(D)$ where D_0 is a compact subset of H with positive measure, and $B = \mathsf{p}_Y(\Omega)$. For $g \in \Omega$ and $x = \mathsf{p}_X(g)$, we have by (4.18) and (4.20),

$$\begin{aligned}
&\left|\pi_X(\Lambda_t)\phi(x) - \pi_X(\Lambda_t^G)\phi(x)\right| \\
&\le |\pi_X(\Lambda_t)\phi(x) - \mathcal{A}_{t-3\varepsilon}(g)| + \left|\mathcal{A}_{t-3\varepsilon}(g) - \pi_X(\Lambda_t^G)\phi(x)\right| \\
&\le |\mathcal{A}_{t+3\varepsilon}(g) - \mathcal{A}_{t-3\varepsilon}(g)| + \left|\mathcal{A}_{t-3\varepsilon}(g) - \pi_X(\Lambda_t^G)\phi(x)\right| \\
&\le |\mathcal{A}_{t+3\varepsilon}(g) - \mathcal{A}_{t-3\varepsilon}(g)| + \left|\mathcal{A}_{t-3\varepsilon}(g) - \mathcal{A}^G_{t-3\varepsilon}(g)\right| + \left|\pi_X(\Lambda_t^G)\phi(x) - \mathcal{A}^G_{t-3\varepsilon}(g)\right| \\
&\le |\mathcal{A}_{t+3\varepsilon}(g) - \mathcal{A}_{t-3\varepsilon}(g)| + \left|\mathcal{A}_{t-3\varepsilon}(g) - \mathcal{A}^G_{t-3\varepsilon}(g)\right| + \left|\mathcal{A}^G_{t+3\varepsilon}(g) - \mathcal{A}^G_{t-3\varepsilon}(g)\right|.
\end{aligned}$$

We will estimate each of these three terms separately.

Recall that $\Omega = \sqcup_{j=1}^m \Omega_j$, and it follows from the definition of $\mathcal{A}_t$ and $\mathcal{A}_t^G$ (see (4.17) and (4.19)) and (CA1)–(CA2) that for $g \in \Omega_j$,

$$\begin{aligned}
&|\mathcal{A}_t(g) - \mathcal{A}_t^G(g)| \\
&= \left|\sum_{i=1}^n \int_{H_t[u_j, \mathsf{r}(v_i)]} F_i(h^{-1}\mathsf{p}_Y(g))\, d\rho(h) - \sum_{i=1}^n \rho(H_t[u_j, \mathsf{r}(v_i)]) \int_Y F_i\, d\mu\right| \\
&\ll \sum_{i=1}^n \rho(H_t) \left|\frac{1}{\rho(H_t[u_j, \mathsf{r}(v_i)])} \int_{H_t[u_j, \mathsf{r}(v_i)]} F_i(h^{-1}\mathsf{p}_Y(g))\, d\rho(h) - \int_Y F_i\, d\mu\right|.
\end{aligned}$$

By Lemma 3.3, we have $F_i \in L^p(B)$. Hence, it follows from the pointwise ergodic theorem in $L^p(B)$ that

$$\lim_{t\to\infty} \frac{|\mathcal{A}_t(g) - \mathcal{A}_t^G(g)|}{\rho(H_t)} = 0 \tag{5.2}$$

for almost every $g \in \Omega$. By (4.20) and Lemma 4.2(i) combined with (A2),

$$\begin{aligned}
\left|\mathcal{A}^G_{t+3\varepsilon}(g) - \mathcal{A}^G_{t-3\varepsilon}(g)\right| &\le \left|\pi_X(\Lambda^G_{t+6\varepsilon})\phi(x) - \pi_X(\Lambda^G_{t-6\varepsilon})\phi(x)\right| \\
&\ll \omega(12\varepsilon)\rho(H_t)\|\phi\|_{L^p(D)}
\end{aligned} \tag{5.3}$$

on a set of full measure of $g \in \Omega$. Since

$$\begin{aligned}
|\mathcal{A}_{t+3\varepsilon}(g) - \mathcal{A}_{t-3\varepsilon}(g)| \le &|\mathcal{A}_{t+3\varepsilon}(g) - \mathcal{A}^G_{t+3\varepsilon}(g)| + |\mathcal{A}^G_{t+3\varepsilon}(g) - \mathcal{A}^G_{t-3\varepsilon}(g)| \\
&+ |\mathcal{A}^G_{t-3\varepsilon}(g) - \mathcal{A}_{t-3\varepsilon}(g)|,
\end{aligned}$$

combining (5.2) and (5.3), we deduce that

$$\limsup_{t\to\infty} \frac{|\mathcal{A}_{t+3\varepsilon}(g) - \mathcal{A}_{t-3\varepsilon}(g)|}{\rho(H_t)} \ll \omega(12\varepsilon)\rho(H_t)\|\phi\|_{L^p(D)} \tag{5.4}$$

on a set of full measure. Finally, we deduce from (5.2), (5.3), and (5.4) that on a set of full measure in D,

$$\limsup_{t\to\infty} \frac{\left|\pi_X(\Lambda_t)\phi(x) - \pi_X(\Lambda_t^G)\phi(x)\right|}{\rho(H_t)} \ll \omega(12\varepsilon)\|\phi\|_{L^p(D)}.$$

Since $\omega(\varepsilon) \to 0$ as $\varepsilon \to 0^+$, this proves the theorem. $\square$

6. Quantitative mean ergodic theorem

6.1. **Admissibility.** The goal of this section is prove the quantitative mean and pointwise ergodic theorems for the normalized sampling operators $\pi_X(\lambda_t)$ defined in (2.9). We shall use notation from Section 2 and assume that the sets G_t satisfy the following additional regularity properties:

(HA1) There exist a basis $\{O_\varepsilon\}_{\varepsilon\in(0,1]}$ of symmetric neighborhoods of the identity in G and $c>0$ such that for every $\varepsilon\in(0,1)$ and $t\ge t_0$

$$O_\varepsilon\cdot G_t\cdot O_\varepsilon\subset G_{t+c\varepsilon}. \tag{6.1}$$

Moreover, for all sufficiently small ε, there exists a nonnegative function $\chi_\varepsilon\in C_c^l(H)$ such that

$$\mathrm{supp}(\chi_\varepsilon)\subset O_\varepsilon,\quad \int_H \chi_\varepsilon\, d\rho=1,\quad \|\chi_\varepsilon\|_{L_l^q(H)}\ll \varepsilon^{-\kappa}, \tag{6.2}$$

and for every compact $\Omega\subset G$ and $\varepsilon\in(0,1)$, there exists a cover

$$\Omega\subset\bigcup_{i=1}^{n_\varepsilon} v_iO_\varepsilon \tag{6.3}$$

with $n_\varepsilon\ll\varepsilon^{-d}$.

(HA2) For every $u\in G$ and for a compact domain $\Omega\subset G$, there exist $c,\theta>0$ such that for every $t\ge t_0$ and $\varepsilon\in(0,1)$,

$$\left(\int_\Omega (\rho(H_{t+\varepsilon}[u,v])-\rho(H_t[u,v]))^r\, dm(v)\right)^{1/r}\le c\,\varepsilon^\theta\rho(H_t).$$

(HA2′) For every compact $\Omega\subset G$, there exist $c,\theta>0$ such that for every $t\ge t_0$ and $\varepsilon\in(0,1)$,

$$\left(\int_{\Omega\times\Omega} (\rho(H_{t+\varepsilon}[u,v])-\rho(H_t[u,v]))^r\, dm(u)dm(v)\right)^{1/r}\le c\,\varepsilon^\theta\rho(H_t).$$

(HS) every $x\in X$ has a neighbourhood U such that there exists a Lipschitz (with respect to the neighbourhoods O_ε) section $\mathsf{s}:U\to G$ of the map p_X.

We recall that when $l>0$ we assume that G is a Lie group and H is a closed subgroup. In this case, we take O_ε to be the symmetric ε-neighbourhoods of identity with respect to a fixed Riemannian metric in G. Then (6.2) holds with $\kappa=l+\dim(H)(1-1/q)$ and (6.3) holds with $d=\dim(G)$. Moreover, in this case (HS) also holds, and one can choose the section s to be smooth on U.

We say that a function $E(t)$ is *coarsely admissible* if there exists $c>0$ such that $\sup_{s\in[t,t+1]}E(s)\le c\,E(t)$ for all t.

Theorem 6.1. *Let $1\le p<q\le\infty$ and $l\in\mathbb{Z}_{\ge0}$. Suppose that* (CA1), (CA2), (HA1), (HA2′) *with $r=pq/(q-p)$, and* (HS) *hold. Assume that for every $g_1,g_2\in G$ and compact $B\subset Y$, the family $\beta_t^{g_1,g_2}$ satisfies the quantitative mean ergodic theorem in $L_l^q(B)^+$ with respect to $\|\cdot\|_{L^p(B)}$, and the error term $E(t)$ is coarsely*

admissible and uniform over g_1, g_2 in a compact subset of G. Namely, for every $F \in L_l^p(B)^+$ and sufficiently large t,

$$\left\| \pi_Y(\beta_t^{g_1,g_2})F - \int_Y F\,d\mu \right\|_{L^p(B)} \ll_{p,q,l,B} E(t)\,\|F\|_{L_l^q(B)}.$$

Then for every compact $D \subset X$, $\phi \in L_l^q(D)^+$ and sufficiently large t,

$$\left\| \pi_X(\lambda_t)\phi - \pi_X(\lambda_t^G)\phi \right\|_{L^p(D)} \ll_{p,q,l,D} E(t)^\delta\,\|\phi\|_{L_l^q(D)}$$

with $\delta > 0$ independent of D and ϕ.

Proof. The proof of theorem will follows the same outline as the proof of Theorem 4.1. Throughout the proof, we shall use a parameter $\varepsilon = \varepsilon(t) \in (0,1)$ such that $\varepsilon(t) \to 0$ as $t \to \infty$, which will be specified later. Because of (HS), decomposing the function ϕ into a sum of functions with small support, we reduce the proof to the situation when there exists a section $\mathsf{s} : D \to G$ of the factor map $\mathsf{p}_X : G \to H\backslash G = X$ such that $\mathsf{s}|_D$ is Lipschitz. Moreover, when G and H are Lie groups, we may assume that s is smooth on D.

Let the function $\chi_\varepsilon \in C_c^l(H)$ be as in (6.2). We define the function $f_\varepsilon : G \to \mathbb{R}$ as in (3.1). Since $\mathsf{s}|_D$ is Lipschitz and D is compact, it follows from (HA1) that there exists $c > 0$ such that for all sufficiently small ε,

$$O_\varepsilon \cdot G_t \cdot g^{-1} O_\varepsilon g \subset G_{t+c\varepsilon} \quad \text{for every } g \in \mathsf{s}(D), \tag{6.4}$$

$$G_t \cdot \mathsf{s}(yO_\varepsilon)^{-1} \subset G_{t+c\varepsilon} \cdot \mathsf{s}(y)^{-1} \quad \text{for every } y \in D. \tag{6.5}$$

Therefore, we may argue as in the proof of Theorem 4.1 (cf. (4.9)–(4.10) and (4.12)–(4.13)) to show that for $u \in G$,

$$\pi_X(\Lambda_t)\phi(\mathsf{p}_X(u)) \le \sum_{\gamma\in\Gamma} \int_{H_{t+c\varepsilon}[u,\mathsf{r}(u\gamma)]} f_\varepsilon(h^{-1}u\gamma)\,d\rho(h), \tag{6.6}$$

$$\pi_X(\Lambda_t)\phi(\mathsf{p}_X(u)) \ge \sum_{\gamma\in\Gamma} \int_{H_{t-c\varepsilon}[u,\mathsf{r}(u\gamma)]} f_\varepsilon(h^{-1}u\gamma)\,d\rho(h), \tag{6.7}$$

and

$$\pi_X(\Lambda_t^G)\phi(\mathsf{p}_X(u)) \le \int_G \int_{H_{t+c\varepsilon}[u,\mathsf{r}(ug)]} f_\varepsilon(h^{-1}ug)\,d\rho(h), \tag{6.8}$$

$$\pi_X(\Lambda_t^G)\phi(\mathsf{p}_X(u)) \ge \int_G \int_{H_{t-c\varepsilon}[u,\mathsf{r}(ug)]} f_\varepsilon(h^{-1}ug)\,d\rho(h)dm(g), \tag{6.9}$$

It follows from (6.5) for every $u \in G$, $g \in H\mathsf{s}(D)$ and $g' \in HgO_\varepsilon$,

$$H_{t-c\varepsilon}[u,\mathsf{r}(g')] \subseteq H_t[u,\mathsf{r}(g)] \subseteq H_{t+c\varepsilon}[u,\mathsf{r}(g')]. \tag{6.10}$$

6.2. **Local analysis.** Let $\Omega := (\overline{O_1 \cap H})^{-1}\mathsf{s}(D)$ and $B = \mathsf{p}_Y(\Omega)$. Then $\mathrm{supp}(f_\varepsilon) \subset \Omega$, and by (HA1), there exists a finite cover

$$\Omega \subset \bigcup_{i=1}^{n_\varepsilon} v_i O_\varepsilon$$

with $v_i \in \Omega$ and $n_\varepsilon \ll \varepsilon^{-d}$. When $l = 0$, we choose a family of bounded measurable functions $\{\psi_{\varepsilon,i}\}_{i=1}^{n_\varepsilon}$ such that $\sum_{i=1}^{n_\varepsilon} \psi_{\varepsilon,i} = 1$ on Ω and $\mathrm{supp}(\psi_{\varepsilon,i}) \subset v_i O_\varepsilon$. When $l > 0$,

we choose a smooth partition of unity $\{\psi_{\varepsilon,i}\}_{i=1}^{n_\varepsilon}$ satisfying the above properties. Using the standard construction of the partition of unity (see, for instance, [R, Th. 2.13]), we get the estimate

$$\|\psi_{\varepsilon,i}\|_{C^l} \ll \varepsilon^{-l}. \tag{6.11}$$

Let $f_{\varepsilon,i} = f_\varepsilon \cdot \psi_{\varepsilon,i}$ and $F_{\varepsilon,i}(\mathsf{p}_Y(g)) = \sum_{\gamma\in\Gamma} f_{\varepsilon,i}(g\gamma)$. For sufficiently small ε, the map $\mathsf{p}_Y : G \to Y = G/\Gamma$ is a bijection on $\mathrm{supp}(f_{\varepsilon,i})$, so that

$$\int_Y F_{\varepsilon,i}\, d\mu = \int_G f_{\varepsilon,i}\, dm \quad \text{and} \quad \|F_{\varepsilon,i}\|_{L_l^q(B)} \ll \|f_{\varepsilon,i}\|_{L_l^q(\Omega)}. \tag{6.12}$$

Using (6.10) we deduce as in the proof of Theorem 4.1 (cf. (4.15)–(4.16)) that

$$\sum_{\gamma\in\Gamma} \int_{H_t[u,\mathsf{r}(u\gamma)]} f_\varepsilon(h^{-1}u\gamma)\, d\rho(h) \le \sum_{i=1}^{n_\varepsilon} \int_{H_{t+c\varepsilon}[u,\mathsf{r}(v_i)]} F_{\varepsilon,i}(h^{-1}\mathsf{p}_Y(u))\, d\rho(h), \tag{6.13}$$

and

$$\sum_{\gamma\in\Gamma} \int_{H_t[u,\mathsf{r}(u\gamma)]} f_\varepsilon(h^{-1}u\gamma)\, d\rho(h) \ge \sum_{i=1}^{n_\varepsilon} \int_{H_{t-c\varepsilon}[u,\mathsf{r}(v_i)]} F_{\varepsilon,i}(h^{-1}\mathsf{p}_Y(u))\, d\rho(h). \tag{6.14}$$

By (HA1) there exist a subset $\{u_j\}_{j=1}^{m_\varepsilon}$ of Ω with $m_\varepsilon \ll \varepsilon^{-d}$ and a measurable partition

$$\Omega = \bigsqcup_{j=1}^{m_\varepsilon} \Omega_j \quad \text{such that} \quad \Omega_j \subset u_j O_\varepsilon.$$

It follows from (6.4) that for all $u \in \Omega_j$,

$$H_{t-c\varepsilon}[u_j, \mathsf{r}(v_i)] \subseteq H_t[u, \mathsf{r}(v_i)] \subseteq H_{t+c\varepsilon}[u_j, \mathsf{r}(v_i)]. \tag{6.15}$$

We introduce a measurable function $\mathcal{A}_t : \Omega \to \mathbb{R}$ defined piecewise by

$$\mathcal{A}_t(u) = \sum_{i=1}^{n_\varepsilon} \int_{H_t[u_j,\mathsf{r}(v_i)]} F_{\varepsilon,i}(h^{-1}\mathsf{p}_Y(u))\, d\rho(h), \quad u \in \Omega_j. \tag{6.16}$$

By (6.15), for $u \in \Omega$,

$$\mathcal{A}_{t-c\varepsilon}(u) \le \sum_{i=1}^{n_\varepsilon} \int_{H_t[u,\mathsf{r}(v_i)]} F_{\varepsilon,i}(h^{-1}\mathsf{p}_Y(u))\, d\rho(h) \le \mathcal{A}_{t+c\varepsilon}(u).$$

Therefore, combining (6.6), (6.7), (6.13), (6.14), we deduce for $u \in \Omega$,

$$\mathcal{A}_{t-3c\varepsilon}(u) \le \pi_X(\Lambda_t)\phi(\mathsf{p}_X(u)) \le \mathcal{A}_{t+3c\varepsilon}(u). \tag{6.17}$$

We also prove similar estimate for the averages $\pi_X(\Lambda_t^G)$. By (6.10) and (6.15), we have for $u \in \Omega_j$,

$$\begin{aligned}\int_G \int_{H_t[u,\mathsf{r}(ug)]} f(h^{-1}ug)\, d\rho(h) dm(g) &\le \sum_{i=1}^{n_\varepsilon} \int_G \int_{H_{t+c\varepsilon}[u,\mathsf{r}(v_i)]} f_{\varepsilon,i}(h^{-1}ug)\, d\rho(h) dm(g) \\ &= \sum_{i=1}^{n} \rho(H_{t+c\varepsilon}[u,\mathsf{r}(v_i)]) \int_G f_{\varepsilon,i}\, dm \\ &\le \sum_{i=1}^{n_\varepsilon} \rho(H_{t+2c\varepsilon}[u_j,\mathsf{r}(v_i)]) \int_G f_{\varepsilon,i}\, dm,\end{aligned}$$

and similarly,

$$\int_G \int_{H_t[u,\mathsf{r}(ug)]} f(h^{-1}ug)\, d\rho(h) dm(g) \ge \sum_{i=1}^{n_\varepsilon} \rho(H_{t-2c\varepsilon}[u_j,\mathsf{r}(v_i)]) \int_G f_{\varepsilon,i}\, dm.$$

Hence, it follows from (6.8) and (6.9) that the function $\mathcal{A}_t^G : \Omega \to \mathbb{R}$ defined by

$$\mathcal{A}_t^G(u) = \sum_{i=1}^{n_\varepsilon} \rho(H_t[u_j,\mathsf{r}(v_i)]) \int_G f_{\varepsilon,i}\, dm, \quad u \in \Omega_j, \tag{6.18}$$

satisfies

$$\mathcal{A}_{t-3c\varepsilon}^G(u) \le \pi_X(\Lambda_t^G)\phi(\mathsf{p}_X(u)) \le \mathcal{A}_{t+3c\varepsilon}^G(u). \tag{6.19}$$

for $u \in \Omega$.

Since by (2.4),

$$\left\|\pi_X(\Lambda_{t+c\varepsilon}^G)\phi \circ \mathsf{p}_X - \pi_X(\Lambda_t^G)\phi \circ \mathsf{p}_X\right\|_{L^p(\Omega)} \ll \left\|\pi_X(\Lambda_{t+c\varepsilon}^G)\phi - \pi_X(\Lambda_t^G)\phi\right\|_{L^p(D)},$$

it follows from Lemma 4.2(ii) combined with (HA2′) that

$$\left\|\pi_X(\Lambda_{t+c\varepsilon}^G)\phi \circ \mathsf{p}_X - \pi_X(\Lambda_t^G)\phi \circ \mathsf{p}_X\right\|_{L^p(\Omega)} \ll \varepsilon^\theta \rho(H_t)\|\phi\|_{L^q(D)}, \tag{6.20}$$

Moreover, combining (6.19) and (6.20), we deduce that

$$\begin{aligned}\left\|\mathcal{A}_{t+c\varepsilon}^G - \mathcal{A}_t^G\right\|_{L^p(\Omega)} &\le \left\|\pi_X(\Lambda_{t+4\varepsilon}^G)\phi \circ \mathsf{p}_X - \pi_X(\Lambda_{t-3\varepsilon}^G)\phi \circ \mathsf{p}_X\right\|_{L^p(\Omega)} \\ &\ll \varepsilon^\theta \rho(H_t)\|\phi\|_{L^q(D)}.\end{aligned} \tag{6.21}$$

6.3. **The duality argument.** Our next task is to estimate $\|\mathcal{A}_t - \mathcal{A}_t^G\|_{L^p(\Omega)}$. Let

$$\begin{aligned}\mathcal{E}_t^{ij}(y) &= \frac{1}{\rho(H_t[u_j,\mathsf{r}(v_i)])} \int_{H_t[u_j,\mathsf{r}(v_i)]} F_{\varepsilon,i}(h^{-1}y)\, d\rho(h) - \int_Y F_{\varepsilon,i}\, d\mu \\ &= \pi_Y(\beta^{u_j,\mathsf{r}(v_i)})F_{\varepsilon,i}(y) - \int_Y F_{\varepsilon,i}\, d\mu.\end{aligned}$$

It follows from the quantitative mean ergodic theorem in $L_l^q(B)^+$ with respect to $\|\cdot\|_{L^p(B)}$ that

$$\|\mathcal{E}_t^{ij}\|_{L^p(B)} \le E(t)\|F_{\varepsilon,i}\|_{L_l^q(B)}.$$

By (6.12) and (6.11),

$$\|F_{\varepsilon,i}\|_{L_l^q(B)} \ll \|f_{\varepsilon,i}\|_{L_l^q(\Omega)} = \|f_\varepsilon \cdot \psi_{\varepsilon,i}\|_{L_l^q(\Omega)} \ll \|\psi_{\varepsilon,i}\|_{C^l} \cdot \|f_\varepsilon\|_{L_l^q(\Omega)} \ll \varepsilon^{-l}\|f_\varepsilon\|_{L_l^q(\Omega)}$$

Recall that when $l>0$, we are assuming that G and H are Lie groups, and the section s is smooth on D. This implies that the map

$$\mathsf{p}_X^{-1}(D)\to H\times D: g\mapsto (\mathsf{h}(g),\mathsf{p}_X(g))$$

is a diffeomorphism. Then it follows from the definition of f_ε (see (3.1)) that

$$\|f_\varepsilon\|_{L^q_l(\Omega)}\ll \|\chi_\varepsilon\|_{L^q_l(H)}\|\phi\|_{L^q_l(D)}\ll \varepsilon^{-\kappa}\|\phi\|_{L^q_l(D)}.$$

A similar estimate when $l=0$ follows from (2.4) and (6.2). Hence, we conclude that

$$\|\mathcal{E}_t^{ij}\|_{L^p(B)}\ll E(t)\varepsilon^{-(l+\kappa)}\|\phi\|_{L^q_l(D)}.$$

Now it follows from the triangle inequality, Lemma 3.2, and (CA1)–(CA2) that

$$\begin{aligned}
&\|\mathcal{A}_t-\mathcal{A}_t^G\|_{L^p(\Omega)}\le \sum_{j=1}^{m_\varepsilon}\|\mathcal{A}_t-\mathcal{A}_t^G\|_{L^p(\Omega_j)}\\
=&\sum_{j=1}^{m_\varepsilon}\left\|\sum_{i=1}^{n_\varepsilon}\int_{H_t[u_j,\mathsf{r}(v_i)]}F_{\varepsilon,i}(h^{-1}\mathsf{p}_Y(g))\,d\rho(h)-\sum_{i=1}^{n_\varepsilon}\rho(H_t[u_j,\mathsf{r}(v_i)])\int_G f_{\varepsilon,i}\,dm\right\|_{L^p(\Omega_j)}\\
\ll&\sum_{i=1}^{n_\varepsilon}\sum_{j=1}^{m_\varepsilon}\rho(H_t[u_j,\mathsf{r}(v_i)])\left\|\mathcal{E}_t^{ij}\right\|_{L^p(B)}\\
\ll&\,\varepsilon^{-2d}\rho(H_t)E(t)\varepsilon^{-(l+\kappa)}\|\phi\|_{L^q_l(D)}.
\end{aligned}\tag{6.22}$$

Combining (6.21) and (6.22), we obtain

$$\begin{aligned}
&\|\mathcal{A}_{t+c\varepsilon}-\mathcal{A}_t\|_{L^p(\Omega)}\\
\le&\|\mathcal{A}_{t+c\varepsilon}-\mathcal{A}_{t+c\varepsilon}^G\|_{L^p(\Omega)}+\|\mathcal{A}_{t+c\varepsilon}^G-\mathcal{A}_t^G\|_{L^p(\Omega)}+\|\mathcal{A}_t^G-\mathcal{A}_t\|_{L^p(\Omega)}\\
\ll&\left(\rho(H_{t+c\varepsilon})E(t+c\varepsilon)+\rho(H_t)E(t)\right)\varepsilon^{-(2d+l+\kappa)}\|\phi\|_{L^q_l(D)}+\varepsilon^\theta\rho(H_t)\|\phi\|_{L^q_l(D)}.
\end{aligned}$$

Hence, it follows from (CA2) and coarse admissibility of the error term $E(t)$ that

$$\|\mathcal{A}_{t+c\varepsilon}-\mathcal{A}_t\|_{L^p(B)}\ll (E(t)\varepsilon^{-(2d+l+\kappa)}+\varepsilon^\theta)\rho(H_t)\|\phi\|_{L^q(D)}.\tag{6.23}$$

By (2.4) and (6.17),

$$\begin{aligned}
&\left\|\pi_X(\Lambda_t)\phi-\pi_X(\Lambda_t^G)\phi\right\|_{L^p(D)}\\
\ll&\left\|\pi_X(\Lambda_t)\phi\circ\mathsf{p}_X-\pi_X(\Lambda_t^G)\phi\circ\mathsf{p}_X\right\|_{L^p(\Omega)}\\
\le&\left\|\pi_X(\Lambda_t)\phi\circ\mathsf{p}_X-\mathcal{A}_{t-3c\varepsilon}\right\|_{L^p(\Omega)}+\left\|\mathcal{A}_{t-3c\varepsilon}-\pi_X(\Lambda_t^G)\phi\circ\mathsf{p}_X\right\|_{L^p(\Omega)}\\
\le&\left\|\mathcal{A}_{t+3c\varepsilon}-\mathcal{A}_{t-3c\varepsilon}\right\|_{L^p(\Omega)}+\left\|\mathcal{A}_{t-3c\varepsilon}-\pi_X(\Lambda_t^G)\phi\circ\mathsf{p}_X\right\|_{L^p(\Omega)},
\end{aligned}$$

and by (6.19),

$$\begin{aligned}
&\left\|\mathcal{A}_{t-3c\varepsilon}-\pi_X(\Lambda_t^G)\phi\circ\mathsf{p}_X\right\|_{L^p(\Omega)}\\
\le&\left\|\mathcal{A}_{t-3c\varepsilon}-\mathcal{A}_{t-3c\varepsilon}^G\right\|_{L^p(\Omega)}+\left\|\pi_X(\Lambda_t^G)\phi\circ\mathsf{p}_X-\mathcal{A}_{t-3c\varepsilon}^G\right\|_{L^p(\Omega)}\\
\le&\left\|\mathcal{A}_{t-3c\varepsilon}-\mathcal{A}_{t-3c\varepsilon}^G\right\|_{L^p(\Omega)}+\left\|\pi_X(\Lambda_t^G)\phi\circ\mathsf{p}_X-\pi_X(\Lambda_{t-6c\varepsilon}^G)\phi\circ\mathsf{p}_X\right\|_{L^p(\Omega)}.
\end{aligned}$$

Therefore, it follows from (6.23), (6.22) combined with coarse admissibility of $E(t)$, and (6.20) that

$$\left\|\pi_X(\Lambda_t)\phi - \pi_X(\Lambda_t^G)\phi\right\|_{L^p(D)} \ll (E(t)\varepsilon^{-(2d+l+\kappa)} + \varepsilon^\theta)\rho(H_t)\|\phi\|_{L_l^q(D)}.$$

This estimate holds for all sufficiently small ε. In order to optimise it, we pick

$$\varepsilon = E(t)^{(\theta+2d+l+\kappa)^{-1}}. \tag{6.24}$$

This proves the theorem with $\delta = \theta/(\theta + 2d + l + \kappa)$. □

7. Quantitative pointwise ergodic theorem

We now turn to the quantitative pointwise ergodic theorem for the normalized sampling operators $\lambda_X(\lambda_t)$.

Theorem 7.1. *Let $1 \le p < q \le \infty$ and $l \in \mathbb{Z}_{\ge 0}$. Suppose that* (CA1), (CA2), (HA1), (HA2) *with $r = q/(q-1)$,* (HA2′) *with $r = pq/(q-p)$, and* (HS) *hold. Assume that for every $g_1, g_2 \in G$ and every compact domain $B \subset Y$, the family $\beta_t^{g_1,g_2}$ satisfies the quantitative mean ergodic theorem in $L_l^q(B)^+$ with respect to $\|\cdot\|_{L^p(B)}$ with exponential rate, and the error term is uniform over g_1, g_2 in a compact subset of G. Then for every compact domain $D \subset X$, a function $\phi \in L_l^p(D)^+$ and almost every $x \in D$, there exists $\delta > 0$ such that*

$$\left|\pi_X(\lambda_t)\phi(x) - \pi_X(\lambda_t^G)\phi(x)\right| \le C(\phi, x)e^{-\delta t}$$

for all $t \ge t_0$. Furthermore, δ is idependent of D, ϕ and x, and $\|C(\phi,\cdot)\|_{L^p(D)} \le C_{p,q}\|\phi\|_{L_l^q(D)}$.

Proof. This proof is a refinement of the proof of Theorem 6.1, and we shall use some of the notation and estimates obtained there. As in that proof, we reduce our argument to the case when the section s is Lipschitz on D. Moreover, if G and H are Lie groups, we reduce the proof to the case when the section s is smooth on D.

By Theorem 6.1, for some $\delta_1 > 0$,

$$\left\|\pi_X(\Lambda_t)\phi - \pi_X(\Lambda_t^G)\phi\right\|_{L^p(D)} \ll e^{-\delta_1 t}\rho(H_t)\|\phi\|_{L_l^q(D)}. \tag{7.1}$$

Let Ω be a compact subset of G as in the proof of Theorem 6.1, and let $\mathcal{A}_t$ and $\mathcal{A}_t^G$ be functions on Ω defined as in (6.16) and (6.18). During the proof of Theorem 6.1, we have established estimate (6.22) with ε as in (6.24) which implies that

$$\left\|\mathcal{A}_t - \mathcal{A}_t^G\right\|_{L^p(\Omega)} \ll e^{-\delta_2 t}\rho(H_t)\|\phi\|_{L_l^q(D)} \tag{7.2}$$

for some $\delta_2 > 0$. Let $\delta = \min\{\delta_1, \delta_2\}$.

We take an increasing sequence $\{t_i\}_{i\ge 0}$ that contains all positive integers greater than t_0 and has uniform spacing $\lfloor e^{p\delta n/4}\rfloor^{-1}$ on the intervals $[n, n+1]$, $n \in \mathbb{N}$. Then

$$t_{i+1} - t_i \ll e^{-p\delta\lfloor t_i\rfloor/4}$$

for all $i \ge 0$. Let

$$C(x,\phi) := \left(\sum_{i\ge 0} e^{p\delta t_i/2}\rho(H_{t_i})^{-p}\left|\pi_X(\Lambda_{t_i})\phi(x) - \pi_X(\Lambda_{t_i}^G)\phi(x)\right|^p\right)^{1/p}.$$

Then we have

$$\left|\pi_X(\Lambda_{t_i})\phi(x) - \pi_X(\Lambda^G_{t_i})\phi(x)\right| \le C(x,\phi) e^{-\delta t_i/2}\rho(H_{t_i}) \tag{7.3}$$

for all $i \ge 0$. It follows from (7.1) that

$$\begin{aligned}\|C(\cdot,\phi)\|^p_{L^p(D)} &= \sum_{i\ge 0} \left\| e^{\delta t_i/2}\rho(H_{t_i})^{-1}\left|\pi_X(\Lambda_{t_i})\phi - \pi_X(\Lambda^G_{t_i})\phi\right|\right\|^p_{L^p(D)} \\ &\ll \sum_{i\ge 0} e^{-p\delta t_i/2}\|\phi\|^p_{L^q_l(D)} \le \sum_{n\ge \lfloor t_0\rfloor} e^{-p\delta n/2}\lfloor e^{p\delta n/4}\rfloor \|\phi\|^p_{L^q_l(D)} < \infty,\end{aligned}$$

and, in particular, $C(x,\phi)$ is finite for almost every $x \in D$.

Using similar argument, we deduce from (7.2) that

$$\left|\mathcal{A}_{t_i}(u) - \mathcal{A}^G_{t_i}(u)\right| \le C'(u,\phi) e^{-\delta t_i/2}\rho(H_{t_i}) \tag{7.4}$$

for all $i \ge 0$, where the estimator $C'(u,\phi)$ is finite for almost all $u \in \Omega$.

To finish the proof, we need to extend estimate (7.3) to general t. We pick $t_i < t$ such that

$$\varepsilon := t - t_i \ll e^{-p\delta\lfloor t_i\rfloor/4} \ll e^{-p\delta t/4}.$$

Let t^+ be the least element of $\{t_i\}_{i\ge 0}$ that satisfies $t_i \ge t + 3c\varepsilon$, and let t^- be the greatest element of $\{t_i\}_{i\ge 0}$ that satisfies $t_i \le t - 3c\varepsilon$. Note that

$$t^+ - t^- \ll e^{-p\delta t/4}.$$

We deduce from (6.17) that for $u \in \Omega$ and $x = \mathsf{p}_X(u) \in D$,

$$\left|\pi_X(\Lambda_t)\phi(x) - \pi_X(\Lambda_{t_i})\phi(x)\right| \le \left|\mathcal{A}_{t^+}(u) - \mathcal{A}_{t^-}(u)\right|.$$

Then it follows from (7.4), (CA2), and (6.21) (where we use again the pointwise estimate arising from summing the differences $e^{\eta t}\left\|\mathcal{A}^G_{t^+} - \mathcal{A}^G_{t^-}\right\|_p$ over the sequence t_i, for suitable η) that for u in a set of full measure in Ω,

$$\begin{aligned}|\mathcal{A}_{t^+}(u) - \mathcal{A}_{t^-}(u)| &\le |\mathcal{A}_{t^+}(u) - \mathcal{A}^G_{t^+}(u)| + |\mathcal{A}^G_{t^+}(u) - \mathcal{A}^G_{t^-}(u)| + |\mathcal{A}^G_{t^-}(u) - \mathcal{A}_{t^-}(u)| \\ &\ll_{\phi,u}\ e^{-\delta t^+/2}\rho(H_{t^+}) + (e^{-p\delta t/4})^\theta \rho(H_{t^- - 3c\varepsilon}) + e^{-\delta t^-/2}\rho(H_{t^-}) \\ &\ll_{\phi,u}\ e^{-\delta' t}\rho(H_t)\end{aligned}$$

with some $\delta' > 0$, and hence for $x = \mathsf{p}_X(u)$,

$$|\pi_X(\Lambda_t)\phi(x) - \pi_X(\Lambda_{t_i})\phi(x)| \ll_{\phi,x} e^{-\delta' t}\rho(H_t). \tag{7.5}$$

Also, it follows from Lemma 4.2(i) combined with (HA2) that for almost every $x \in X$,

$$\left|\pi_X(\Lambda^G_t)\phi(x) - \pi_X(\Lambda^G_{t_i})\phi(x)\right| \ll_{\phi,x} (e^{-p\delta t/4})^\theta \rho(H_t). \tag{7.6}$$

Finally, combining (7.5), (7.3), (7.6) we conclude that for x in a set of full measure in D,

$$\begin{aligned}\left|\pi_X(\Lambda_t)\phi(x) - \pi_X(\Lambda^G_t)\phi(x)\right| &\le |\pi_X(\Lambda_t)\phi(x) - \pi_X(\Lambda_{t_i})\phi(x)| \\ &\quad + \left|\pi_X(\Lambda_{t_i})\phi(x) - \pi_X(\Lambda^G_{t_i})\phi(x)\right| \\ &\quad + \left|\pi_X(\Lambda^G_t)\phi(x) - \pi_X(\Lambda^G_{t_i})\phi(x)\right| \\ &\ll_{\phi,x} e^{-\delta'' t}\rho(H_t)\end{aligned}$$

with some $\delta'' > 0$.

It is clear from the foregoing proof that as x varies over the compact domain D, the constant $C(\phi, x)$ implied in the last estimate satisfies the integrability properties stated in the theorem, namely $\|C(\phi,\cdot)\|_{L^p(D)} \le C_{p,q}\|\phi\|_{L^q_l(D)}$. This completes the proof of Theorem 7.1. □

8. Volume estimates

8.1. **Volume asymptotics on algebraic varieties.** Let $G \subset \mathrm{SL}_d(\mathbb{R})$ be a real almost algebraic group, and H a noncompact almost algebraic subgroup of G. Let ρ denote a left Haar measure on H and m a left Haar measure on G. We fix a non-negative proper homogeneous polynomial P on $\mathrm{Mat}_d(\mathbb{R})$, and consider the family of sets

$$H_t := \{h \in H : \log P(h) \le t\}. \tag{8.1}$$

The aim of this section is to discuss the properties of the sets H_t and their volumes. In order to prove our main results stated in Section 1, we will also need to consider more general families of sets defined by

$$H_t[g_1, g_2] := \{h \in H : \log P(g_1^{-1} h g_2) \le t\} \tag{8.2}$$

for $g_1, g_2 \in G$.

We recall the results [GN1, Th.7.17–7.18]. While these results were stated for algebraic sets, the proof, which is based on resolution of singularities, applies to semialgebraic sets as well and, in particular, to the almost algebraic group H.

Theorem 8.1 ([GN1])**.** (i) (volume asymptotics) *There exist $a \in \mathbb{Q}_{\ge 0}$, $b \in \mathbb{Z}_{\ge 0}$, and $\delta_0 > 0$ such that*

$$\rho(H_t) = e^{at}\left(\sum_{i=0}^{b} c_i t^i\right) + O\left(e^{(a-\delta_0)t}\right)$$

for all $t \ge t_0$, where $c_b > 0$.

(ii) (volume regularity) *There exist $c, \theta > 0$ such that the estimate*

$$\rho(H_{t+\varepsilon}) - \rho(H_t) \le c\,\varepsilon^\theta \rho(H_t) \tag{8.3}$$

holds for all $t \ge t_0$ and $\varepsilon \in (0,1)$.

We note that since we are assuming that H is noncompact, it follows that $\rho(H_t) \to \infty$ as $t \to \infty$, and hence $(a, b) \ne (0, 0)$.

8.2. **Polynomial volume growth of the restricted sets.** We say that the group H is of *exponential type* if $a > 0$, and of *subexponential type* otherwise. The following lemma gives a group-theoretic characterisation of these notions. In particular, it implies that they do not depend on a choice of the polynomial P, and on a choice of the embedding of H in $\mathrm{SL}_d(\mathbb{R})$.

Lemma 8.2. *The group H is of subexponential type if and only if its Zariski closure is an almost direct product of a compact subgroup and an abelian $\mathbb{R}$-diagonalisable subgroup.*

Proof. Let $\|\cdot\|$ denote the Euclidean norm on $\mathrm{Mat}_d(\mathbb{R})$. Then since P is non-negative proper and homogeneous of degree $\deg P$, it follows by compactness of the unit sphere that there exists $C > 1$ such that

$$C^{-1}\,\|x\|^{\deg(P)} \le P(x) \le C\,\|x\|^{\deg(P)} \tag{8.4}$$

for all $x \in \mathrm{Mat}_d(\mathbb{R})$. Therefore, it is sufficient to prove the claim for the sets $H_t = \{h \in H :\, \log\|h\| \le t\}$.

Since H has finite index in its Zariski closure, it has exponential type if and only if the algebraic envelop does. Hence, we may assume that H is algebraic. The group H has the Levi decomposition $H = LAU$, where L is a semisimple almost algebraic subgroup, A is an abelian $\mathbb{R}$-diagonalizable almost algebraic subgroup, and U is a unipotent normal algebraic subgroup. Moreover, LA is an almost direct product of L and A. Under the product map, the left invariant measure ρ on H is equal (up to a constant factor) to the product of the invariant measures on the (unimodular) factors. If Q is a compact subset of LA of positive measure, then there exists $c > 0$ such that $\log\|kh\| \le \log\|h\| + c$ for all $k \in Q$ and $h \in H$. Hence,

$$QU_t \subset H_{t+c} \quad \text{and} \quad \rho(H_{t+c}) \gg \mathrm{vol}(U_t).$$

This implies that if H is not of exponential type, then U is not of exponential type as well. Suppose that $U \ne 1$. The exponential map $\exp : \mathrm{Lie}(U) \to U$ is a diffeomorphism and the invariant measure on U is up to a constant equal to the image under exp of the Lebesgue measure on $\mathrm{Lie}(U)$. Since there exists $c > 0$ such that $\|\exp(v)\| \le c\|v\|^d$ for all $v \in \mathrm{Lie}(U)$, it follows that

$$\mathrm{vol}(U_t) \ge \mathrm{vol}(\{v \in \mathrm{Lie}(U) :\, \|v\| \le (c^{-1}\,e^t)^{1/d}\}).$$

This gives a contradiction and shows that $U = 1$.

If L is not compact, then it is of exponential type as follows from (see [GW, Sec. 7] and [Mau]). Since for a compact $Q \subset A$, there exists $c > 0$ such that $L_tQ \subset H_{t+c}$. As above, this would imply that H is of exponential type. Hence, we conclude that L must be compact, which completes the proof of the lemma. □

8.3. **The limiting density in the ergodic theorem.** Let O_ε denotes the symmetric neighborhood of identity in G with respect to a Riemannian metric, and $G_t := \{g \in G :\, \log P(g) \le t\}$.

Lemma 8.3. (i) *Given a compact subset Ω of G, there exists $c = c(\Omega) > 0$, such that for every $t \ge t_0$*

$$\Omega \cdot G_t \cdot \Omega \subset G_{t+c}.$$

(ii) *There exists $c > 0$ such that for every $\varepsilon \in (0,1)$ and $t \ge t_0$,*

$$O_\varepsilon \cdot G_t \cdot O_\varepsilon \subset G_{t+c\varepsilon}.$$

Proof. In order to prove (i), it is sufficient to show that there exists $C > 0$ such that for every $b_1, b_2 \in \Omega$ and $x \in \mathrm{Mat}_d(\mathbb{R})$, we have

$$P(b_1 x b_2) \le C\,P(x).$$

Since Ω is compact,

$$P(b_1xb_2) \ll_\Omega \left(\max_{i,j}|x_{ij}|\right)^{\deg(P)}, \quad x \in \mathrm{Mat}_d(\mathbb{R}),$$

and since P is non-negative, proper and homogeneous, it follows by compactness that

$$\left(\max_{i,j}|x_{ij}|\right)^{\deg(P)} \ll P(x), \quad x \in \mathrm{Mat}_d(\mathbb{R}). \tag{8.5}$$

This completes the proof of (i).

To prove (ii), we observe that for every $b_1, b_2 \in O_\varepsilon$ and $x \in \mathrm{Mat}_d(\mathbb{R})$,

$$P(b_1xb_2) - P(x) \ll \varepsilon \left(\max_{i,j}|x_{ij}|\right)^{\deg(P)} \ll \varepsilon\, P(x).$$

This implies (ii). □

The following lemma follows immediately from Lemma 8.3.

Lemma 8.4. (i) *Given a compact subset Ω of G, there exists $c = c(\Omega) > 0$, such that for every $g_1, g_2 \in G$, $b_1, b_2 \in \Omega$, and $t \geq t_0$*

$$H_{t-c}[g_1, g_2] \subset H_t[g_1b_1, g_2b_2] \subset H_{t+c}[g_1, g_2].$$

(ii) *There exists $c > 0$ such that for every $g_1, g_2 \in G$, $b_1, b_2 \in O_\varepsilon$ with $\varepsilon \in (0, 1)$, and $t \geq t_0$,*

$$H_{t-c\varepsilon}[g_1, g_2] \subset H_t[g_1b_1, g_2b_2] \subset H_{t+c\varepsilon}[g_1, g_2].$$

The next proposition justifies existence of the limit measures ν_x defined in (2.7).

Proposition 8.5. *For every $g_1, g_2 \in G$, the limit*

$$\Theta(g_1, g_2) := \lim_{t\to\infty} \frac{\rho(H_t[g_1, g_2])}{\rho(H_t)}$$

exists. Moreover, the function Θ is positive and continuous.

Proof. It follows from Theorem 8.1(i) applied to the homogeneous polynomial $P(g_1^{-1}xg_2)$ that

$$\rho(H_t[g_1, g_2]) \sim c(g_1, g_2)e^{a(g_1,g_2)t}t^{b(g_1,g_2)} \quad \text{as } t \to \infty, \tag{8.6}$$

for some $c(g_1, g_2) > 0$, $a(g_1, g_2) \in \mathbb{Q}_{\geq 0}$, $b(g_1, g_2) \in \mathbb{Z}_{\geq 0}$. Moreover, it follows from Lemma 8.4(i) that $a(g_1, g_2)$ and $b(g_1, g_2)$ are independent of $g_1, g_2 \in G$. This implies that the limit exists and is positive.

To prove continuity, we observe that for every $b_1, b_2 \in O_\varepsilon$,

$$|\Theta(g_1b_1, g_2b_2) - \Theta(g_1, g_2)| = \lim_{t\to\infty} \frac{|\rho(H_t[g_1b_1, g_2b_2]) - \rho(H_t[g_1, g_2])|}{\rho(H_t)}$$

$$= \max\left\{ \lim_{t\to\infty} \frac{\rho(H_t[g_1b_1, g_2b_2]) - \rho(H_t[g_1, g_2])}{\rho(H_t)}, \lim_{t\to\infty} \frac{\rho(H_t[g_1, g_2]) - \rho(H_t[g_1b_1, g_2b_2])}{\rho(H_t)} \right\},$$

By Lemma 8.4(ii) and (8.6),

$$\begin{aligned}\lim_{t\to\infty} \frac{\rho(H_t[g_1b_1, g_2b_2]) - \rho(H_t[g_1, g_2])}{\rho(H_t)} &\le \lim_{t\to\infty} \frac{\rho(H_{t+c\varepsilon}[g_1, g_2]) - \rho(H_t[g_1, g_2])}{\rho(H_t)} \\ &= \frac{c(g_1, g_2)}{c(e, e)} \left(e^{a(c\varepsilon)} - 1\right) \ll \varepsilon.\end{aligned}$$

Similarly,

$$\lim_{t\to\infty} \frac{\rho(H_t[g_1, g_2]) - \rho(H_t[g_1b_1, g_2b_2])}{\rho(H_t)} \le \frac{c(g_1, g_2)}{c(e, e)} \left(1 - e^{-a(c\varepsilon)}\right) \ll \varepsilon.$$

This proves continuity. □

In the case of groups of subexponential type, we have the following asymptotic formula for the measure of the sets $H_t[g_1, g_2]$, which is independent of g_1, g_2, generalizing Theorem 8.1(i).

Proposition 8.6. *Let $H \subset G$ be of subexponential type. Then there exists $c_b > 0$ and $b \in \mathbb{N}$ such that uniformly over g_1, g_2 in compact subsets of G,*

$$\rho(H_t[g_1, g_2]) = c_b\, t^b + O(t^{b-1})$$

for all $t \ge t_0$.

Proof. Let $\|\cdot\|$ be a Euclidean norm on $\mathrm{Mat}_d(\mathbb{R})$, and $H'_t[g_1, g_2]$ the corresponding balls in H. It follows from (8.4) that there exists $c > 0$ such that for every $g_1, g_2 \in G$ and $t \ge t_0$,

$$H'_{t-c}[g_1, g_2] \subset H_t[g_1, g_2] \subset H'_{t+c}[g_1, g_2].$$

Therefore, it is sufficient to prove the claim of the lemma for a Euclidean norm.

By Lemma 8.4(i), there exists $c > 0$, uniform over g_1, g_2 in compact sets, such that for every $t \ge t_0$,

$$H_{t-c} \subset H_t[g_1, g_2] \subset H_{t+c}.$$

Since by Theorem 8.1(i), we have

$$\rho(H_t) = c_b\, t^b + O(t^{b-1}),$$

this implies the proposition. □

Remark 8.7. Proposition 8.6 implies that for the groups of subexponential type, the regularity properties of sets $H_t[g_1, g_2]$ are straightforward to establish. In particular, the function Θ in Proposition 8.5 is constant, and the claim of Proposition 8.8 below follows directly from Proposition 8.6.

We also note that the argument of Proposition 8.6 applies to sets defined by $P(x) = \|x\|$ where $\|\cdot\|$ is a general norm on $\mathrm{Mat}_d(\mathbb{R})$, not necessarily a *polynomial* one. This implies that Theorem 1.1 holds for the averages along the sets $\Gamma_t := \{\gamma \in \Gamma : \log\|\gamma\| \le t\}$ defined by general norms.

8.4. **Volume regularity properties : average admissibility.** The following proposition gives an averaged version of Theorem 8.1(ii).

Proposition 8.8. *Let $1 \le r < \infty$ and Ω be a compact subset of G.*

(i) *For every $u \in G$, there exist $c = c(u, \Omega, r) > 0$ and $\theta = \theta(u, \Omega, r) > 0$ such that the estimate*

$$\left(\int_\Omega (\rho(H_{t+\varepsilon}[u,v]) - \rho(H_t[u,v]))^r \, dm(v)\right)^{1/r} \le c\,\varepsilon^\theta \rho(H_t)$$

holds for all $t \ge t_0$ and $\varepsilon \in (0,1)$.

(ii) *There exist $c = c(\Omega, r) > 0$ and $\theta = \theta(\Omega, r) > 0$ such that the estimate*

$$\left(\int_{\Omega\times\Omega} (\rho(H_{t+\varepsilon}[u,v]) - \rho(H_t[u,v]))^r \, dm(u)dm(v)\right)^{1/r} \le c\,\varepsilon^\theta \rho(H_t)$$

holds for all $t \ge t_0$ and $\varepsilon \in (0,1)$.

Proof. Since the proofs of (i) and (ii) is very similar, we only present the proof of (i).

Without loss of generality, we may assume that Ω is bounded semialgebraic set. We first prove the assertion when $r = n$ is an integer. Let

$$v(t) := \int_\Omega \rho(H_t[u,v])^n \, dm(v).$$

We claim that for some $c_1, \theta_1 > 0$ this function satisfies the estimate

$$v(t+\varepsilon) - v(t) \le c_1\,\varepsilon^{\theta_1}\, v(t) \tag{8.7}$$

for all $t \ge t_0$ and $\varepsilon \in (0,1)$. We note that by Lemma 8.4(i) and Theorem 8.1(i), $v(t) \ll \rho(H_t)^n$. Therefore, using the inequality $(a-b)^n \le a^n - b^n$ with $a \ge b \ge 0$, we conclude that (8.7) implies (ii) with $p = n$.

Now to prove (8.7), we observe that

$$\begin{aligned} v(t) &= \int_{\Omega\times H^n} \chi_{\{\log P(u^{-1}h_1 v)\le t, \cdots, \log P(u^{-1}h_n v)\le t\}} \, dm(v)d\rho(h_1)\cdots d\rho(h_n) \\ &= \int_{\Omega\times H^n} \chi_{\{\log \Psi(u,v,h_1,\ldots,h_n)\le t\}} \, dm(v)d\rho(h_1)\ldots d\rho(h_n), \end{aligned}$$

where $\Psi(v, h_1, \ldots, h_n) = \max\{P(u^{-1}h_1 v), \ldots, P(u^{-1}h_n v)\}$ is a semialgebraic function on $\Omega \times H^n$. Let $\overline{H}$ denote the projective closure of H. Then Ψ^{-1} is a semialgebraic function on $\Omega \times \overline{H}^n$ that vanishes on the complement of $\Omega \times H^n$. Now we can apply the argument of [GN1, Theorems 7.17] to deduce (8.7).

To prove (i) for general $r \ge 1$, we observe that Hölder's inequality with $q = (\lfloor r \rfloor + 1)/r$ gives

$$\begin{aligned} &\int_\Omega (\rho(H_{t+\varepsilon}[u,v]) - \rho(H_t[u,v]))^r \, dm(v) \\ &\le m(\Omega)^{1-1/q} \left(\int_{\Omega\times\Omega} (\rho(H_{t+\varepsilon}[u,v]) - \rho(H_t[u,v]))^{rq} \, dm(v)\right)^{1/q}. \end{aligned}$$

Hence, the general estimate follows from the case when r is an integer. □

Remark 8.9. Let us note that the quality of the estimate stated in Proposition 8.8(i) is a key ingredient controlling the quality of the mean and pointwise ergodic theorems. Obtaining results of the quality stated in Theorem 1.4 hinges upon establishing an estimate which is uniform in the rate θ and the constant c as u varies in compact sets in G. Since $\mathsf{s}(x)$ and $\mathsf{s}(y)$ are not in H, the integral in Proposition 8.8(i) depends non-trivially on $\mathsf{s}(x)$ and $\mathsf{s}(y)$, so there is no obvious way to exploit invariance of the measure. Rather, the proofs of Proposition 8.8(i) and of Proposition 8.10 below apply resolution of singularities to the parametric family of polynomial maps $h \mapsto P(\mathsf{s}(x)^{-1}h\mathsf{s}(y))$. The estimate produced as a result of this procedure for a given polynomial in the family depends on its coefficients, which in turn depend non-trivially on $\mathsf{s}(x)$ and $\mathsf{s}(y)$. We will establish uniform estimates for the parametric family of polynomials that arises when P is a norm, and this accounts for the appearance of this assumption in Theorem 1.4.

8.5. **Volume regularity properties : subanalytic functions.** The following proposition refines Theorem 8.1(i) and generalizes the discussion to the case of a general subanalytic function.

Proposition 8.10. *Let $a \in \mathbb{Q}_{\geq 0}$ and $b \in \mathbb{Z}_{\geq 0}$ be as in Theorem 8.1(i). Then for every nonnegative continuous subanalytic function $\phi \neq 0$ with compact support and $x \in X$, there exists $\delta > 0$ such that for all $t \geq t_0$,*

$$\int_{G_t} \phi(xg)\, dm(g) = e^{at}\left(\sum_{i=0}^{b} c_i(\phi, x)t^i\right) + O_{\phi,x}\left(e^{(a-\delta)t}\right), \tag{8.8}$$

where $c_b(\phi, x) > 0$.

Proof. Decomposing ϕ into a sum of subanalytic functions with compact supports, we reduce the proof to the case when $\mathrm{supp}(\phi)$ is contained a compact semianalytic set D, and there exists a section $\mathsf{s} : X \to G$ of the factor map $\mathsf{p}_X : G \to H\backslash G = X$ such that $\mathsf{s}|_D$ is analytic. It follows from invariance of m and (2.4) that

$$\begin{aligned} v(t) := \int_{G_t} \phi(xg)\, dm(g) &= \int_{G_t} \phi(\mathsf{p}_X(\mathsf{s}(x)g))\, dm(g) \\ &= \int_{(y,h):\, \mathsf{s}(x)^{-1}h\mathsf{s}(y)\in G_t} \phi(\mathsf{p}_X(h\mathsf{s}(y)))\, d\rho(h)d\xi(y) \\ &= \int_D \phi(y)\rho(H_t[\mathsf{s}(x), \mathsf{s}(y)])\, d\xi(y). \end{aligned}$$

We also note that it follows from (2.4) that the measure ξ is given by an analytic differential form on D. We consider the transform of $v(t)$:

$$f(s) = \int_0^\infty t^{-s} v(\log t) dt.$$

Note that it follows from Lemma 8.4(i) and Theorem 8.1(i) that for some $c > 0$,

$$v(t) \leq \|\phi\|_\infty \xi(D)\rho(H_{t+c}) \ll e^{at}t^b \tag{8.9}$$

and, in particular, the integral $f(s)$ converges when $\mathrm{Re}(s)$ is sufficiently large. In this region, we have

$$\begin{aligned} f(s) &= \int_D \phi(y) \left(\int_0^\infty t^{-s} \rho(H_{\log t}[\mathsf{s}(x), \mathsf{s}(y)]) dt \right) d\xi(y) \\ &= (s-1)^{-1} \int_D \phi(y) \left(\int_H P(\mathsf{s}(x)^{-1} h \mathsf{s}(y))^{-s+1} d\rho(h) \right) d\xi(y) \\ &= (s-1)^{-1} \int_{D\times H} \phi(y) P(\mathsf{s}(x)^{-1} h \mathsf{s}(y))^{-s+1} d\xi(y) d\rho(h). \end{aligned}$$

We observe that the map $(y \times h) \mapsto P(\mathsf{s}(x)^{-1} h \mathsf{s}(y))^{-1}$ extends to a semianalytic function $D \times \overline{H}$, where $\overline{H}$ denotes the projective closure of H, and vanishes on $D \times (\overline{H} - H)$. Now to finish the proof, we can apply the argument of [GN1, Theorems 7.17], but instead of the Hironaka resolution of singularities, we use the rectilinearization of subanalytic functions [P, Theorem 2.7]. Hence, we conclude that (8.8) holds, but a priori the parameters a and b in (8.8) may depend on ϕ and x. However, we observe that since ϕ is continuous, there exists bounded open $O \subset X$ and $m_0 > 0$ such that $\phi(y) \ge m_0$ for all $y \in O$. Therefore, we deduce from Lemma 8.4(i) and Theorem 8.1(i) that for some $c > 0$,

$$v(t) = \int_X \phi(y) \rho(H_t[\mathsf{s}(x), \mathsf{s}(y)])\, d\xi(y) \ge m_0 \xi(O) \rho(H_{t-c}) \gg e^{at} t^b.$$

Combining this estimate with (8.9), we conclude that a and b are independent of ϕ and x. $\square$

Remark 8.11. Let $G = \prod_{i=1}^l G_i$ where $G_i \subset \mathrm{SL}_{d_i}(\mathbb{R})$ be a real almost algebraic group, and let P_i's be non-negative proper homogeneous polynomials on $\mathrm{Mat}_{d_i}(\mathbb{R})$. For a non-compact real almost algebraic subgroup H of G, we consider the sets

$$H_t := \{h \in H :\, \log(P_1(h_1) \cdots P_l(h_l)) \le t\},$$

which appear in number-theoretic applications as height function Then the arguments developed in this section apply to such sets, and the theorems established in the Introduction hold for averages supported on the sets

$$\Gamma_t := \{\gamma \in \Gamma :\, \log(P_1(\gamma_1) \cdots P_l(\gamma_l)) \le t\}.$$

9. Ergodic theory of algebraic subgroups

As noted in the introduction, the possibility of applying the duality principle to establish ergodic theorems for (properly normalized) sampling operators for Γ acting on $H \setminus G$ depends on the validity of ergodic theorems for averages on H acting on G/Γ. We therefore turn now to consider the ergodic theory of algebraic subgroups, namely to consider an algebraic group G acting by measure-preserving transformation on a probability space Y, and to the study of the action restricted to an algebraic subgroup H. In the discussion of this problem it is natural to consider spaces Y more general than just G/Γ (so that G is no longer transitive), and also general closed subgroups of G which are not necessarily algebraic.

We will first consider in §9.1 the case where the volume growth of the restricted sets H_t is subexponential, where one can apply the traditional arguments regarding

regular Følner families to obtain ergodic theorems for actions on a general probability space Y.

We will then assume that H is a subgroup of a connected semisimple Lie group G, which acts on a probability space Y with a strong spectral gap, namely such that each of the simple components of G has a spectral gap. Under this assumption we will prove quantitative mean, maximal and pointwise ergodic theorems, of two kinds. In §9.2 and §9.3 we assume that Y is a manifold and establish quantitative mean ergodic theorem in Sobolev spaces, for a general closed subgroup, including of course connected algebraic subgroups H. When the volume growth of H_t is exponential, the rate of convergence in the mean ergodic theorem will be exponentially fast, and we also establish then an exponentially fast pointwise ergodic theorem for bounded functions.

In §9.4 and §9.5 we turn from Sobolev spaces to Lebesgue spaces. First, in §9.4 we will establish spectral estimates for ergodic averages in actions of general non-amenable closed subgroups which are non-amenably embedded — a term we will define and explain there. In §9.5 we use these estimates and establish exponentially fast maximal, mean and pointwise ergodic theorems for these averages acting in Lebesgue spaces. The proof will in fact only depend on a mild regularity assumption on the averages, an assumption that is always satisfied when the subgroup is algebraic and the averages are defined by a homogeneous proper non-negative polynomial.

Let $H \subset G \subset \mathrm{SL}_d(\mathbb{R})$ be closed noncompact subgroup with a left Haar measure ρ. For an arbitrary measure-preserving action of H on a probability space (Y, μ), we consider the family of averaging operators $\pi_Y(\beta_t) : L^p(Y) \to L^p(Y)$ defined by

$$\pi_Y(\beta_t)F(y) = \frac{1}{\rho(H_t)} \int_{H_t} F(h^{-1}y)\, d\rho(h), \quad F \in L^p(Y),$$

where H_t are the balls associated with a proper non-negative homogeneous polynomial, as defined in Section 8. We will also consider the sets $H_t[g_1, g_2] \subset H$, defined via the embedding $H \subset G$ (see (2.6)).

9.1. **Ergodic theorems in the presence of subexponential growth.**

Theorem 9.1. *Keeping the notation of the preceding paragraph, assume that H and G are almost algebraic, the sets $H_t \subset H$ have subexponential volume growth and that H acts on an arbitrary probability measure space (Y, μ) preserving the measure. Then the averages $\pi_Y(\beta_t)$ satisfy the following.*

(i) Weak-type $(1,1)$-maximal inequality. *For every $F \in L^1(Y)$ and $\delta > 0$,*

$$\mu\left(\left\{\sup_{t \ge t_0} |\pi_Y(\beta_t)F| > \delta\right\}\right) \ll \frac{\|F\|_{L^1(Y)}}{\delta}.$$

(ii) Strong maximal inequality. *For $1 < p \le \infty$ and every $F \in L^p(Y)$,*

$$\left\|\sup_{t \ge t_0} |\pi_Y(\beta_t)F|\right\|_{L^p(Y)} \ll_p \|F\|_{L^p(Y)}.$$

(iii) Mean and pointwise ergodic theorem. *For every* $1 \le p < \infty$ *and* $F \in L^p(Y)$, *the averages* $\pi_Y(\beta_t)F$ *converges almost everywhere and in* L^p*-norm as* $t \to \infty$. *In the ergodic case, the limit is* $\int_Y F d\mu$.

The same results hold without change for each of the families $\pi_Y(\beta_t^{g_1,g_2})$ *supported on* $H_t[g_1, g_2]$, *for any* $g_1, g_2 \in G$.

Before starting the proof, let us observe that by Lemma 8.2, whenever the sets H_t have subexponential volume growth, the group H is in fact amenable, so that we can use classical methods from the ergodic theory of amenable groups (see, for instance, [Ne] for a recent survey, and [AAB] for a detailed discussion). We say that a family of subsets $\{B_t\}$ of H is *asymptotically invariant, or uniform Følner* if for every compact subset Q of H,

$$\lim_{t\to\infty} \frac{\rho(QB_t \triangle B_t)}{\rho(B_t)} = 0. \tag{9.1}$$

We say that a family $\{B_t\}$ is *regular* if

$$\rho(B_t \cdot B_t^{-1}) \ll \rho(B_t). \tag{9.2}$$

Proposition 9.2. *Assume that the sets* $H_t \subset H$ *have subexponential volume growth. Then the family* $\{H_t\}$ *is uniform Følner and regular.*

Proof. Since H is of subexponential type, by Theorem 8.1(i),

$$\rho(H_t) \sim c_b\, t^b \quad \text{ as } t \to \infty \tag{9.3}$$

for some $c_b > 0$ and $b \in \mathbb{N}$.

Without loss of generality, we may assume that the compact set Q in (9.1) contains the identity. There exists $c > 0$ such that $H_tQ \subset H_{t+c}$, so that

$$\limsup_{t\to\infty} \frac{\rho(QH_t \triangle H_t)}{\rho(H_t)} \le \limsup_{t\to\infty} \frac{\rho(H_{t+c} - H_t)}{\rho(H_t)} = 0$$

by (9.3). This proves that $\{H_t\}$ is uniform Følner. Clearly by property (CA2) namely coarse admissibility of H_t, the families $H_t[g_1, g_2]$ are also Følner.

To prove the second claim, we observe that by for every $h, h' \in H$,

$$P(h \cdot h') \ll \left(\max_{ij} |h_{ij}|\right)^{\deg(P)} \left(\max_{ij} |h'_{ij}|\right)^{\deg(P)} \ll P(h)P(h'),$$

and

$$P(h^{-1}) \ll \left(\max_{ij} |h_{ij}|\right)^{(d-1)\deg(P)} \ll P(h)^{d-1}.$$

Therefore, there exists $c > 0$ such that

$$H_t \cdot H_t^{-1} \subset H_{dt+\log(2c)}.$$

Hence, it follows from (9.3) that $\{H_t\}$ is a regular Følner family, and similarly the same holds for each family $H_t[g_1, g_2]$. □

Proof of Theorem 9.1. Since the family of sets H_t is uniform Følner and regular. The theorem is a partial case of [Ne, Th. 6.6]. □

9.2. **Quantitative mean ergodic theorem in Sobolev spaces.** In the present subsection we turn to establish a quantitative mean ergodic theorem for the operators $\pi_Y(\beta_t)$ in Sobolev spaces. This result will be used §9.3 below in the proof of the quantitative pointwise ergodic theorem in Sobolev spaces. We begin by considering a general closed subgroup H, contained in a semisimple Lie group G acting smoothly on a manifold Y with a strong spectral gap, namely each of the simple factors of G has a spectral gap. This condition is of course necessary in order to obtain norm decay along subgroups.

Theorem 9.3. Quantitative mean ergodic theorem in Sobolev spaces. *Assume that*

- *the group H is an arbitrary closed subgroup of a connected semisimple Lie group G with finite centre, and G acts on a manifold Y preserving a probability measure μ,*
- *the representation of every simple factor of G on $L_0^2(Y)$ is isolated from the trivial representation.*

Then there exist $l \in \mathbb{N}$ and $t_0 > 0$ such that for every $1 < p < \infty$, a compact domain B of Y, and $F \in L_l^p(B)$, the following estimate holds with $\kappa_p > 0$, for all $t \geq t_0$,

$$\left\| \pi_Y(\beta_t)F - \int_Y F\, d\mu \right\|_{L^p(Y)} \ll_{p,B} \rho(H_t)^{-\kappa_p} \|F\|_{L_l^p(B)}$$

Furthermore, the rate of convergence applies to each famiy $\pi_Y(\beta_t^{g_1,g_2})$ supported on the sets $H_t[g_1, g_2]$, uniformly when $g_1, g_2 \in G$ vary in compact sets in G.

We remark that the mean ergodic Theorem 9.3 above, in the case where the subgroup H is either amenable, or non-amenable but amenably embedded in G, is the best possible result of its kind. Indeed, in these cases, while quantitative mean ergodic theorems will presently be shown to hold in Sobolev spaces, they definitely do not hold in Lebesgue spaces : the norm of the operators $\pi_Y^0(\beta_t)|_{L^2(Y)\to L^2(Y)}$ is either identically 1 (when H is amenable) or converges to 1 (when H is amenably embedded).

As noted already, under the assumption that H is non-amenable and not embedded amenably, we will derive in Theorem 9.12 in §9.5 below the stronger exponentially fast mean and pointwise ergodic theorems in Lebesgue space L^p, $1 < p < \infty$, going beyond the results in Sobolev spaces.

Proof. We first consider the case when $p = 2$. By [KM1, 2.4.3], there exist $l \in \mathbb{N}$ and $\kappa > 0$ such that for every $F_1, F_2 \in L_l^2(B)$ with zero integrals, we have

$$|\langle \pi_Y(g)F_1, F_2\rangle| \ll \|g\|^{-\kappa} \|F_1\|_{2,l} \|F_2\|_{2,l},$$

where $\|\cdot\|_{2,l}$ denotes the Sobolev norm as defined in [KM1]. While these Sobolev norms are different from the Sobolev norms $\|\cdot\|_{L_l^2(B)}$ that we use in our paper, it is clear that $\|\cdot\|_{2,l} \ll_B \|\cdot\|_{L_l^2(B)}$ on $L_l^2(B)$.

It follows from the above estimate that for every $F \in L_l^2(B)$ with zero integral,

$$\begin{aligned}
\|\pi_Y(\beta_t)F\|^2_{L^2(Y)} =& \frac{1}{\rho(H_t)^2}\int_{H_t\times H_t} \langle \pi_Y(h_1)F, \pi_Y(h_2)F\rangle \, d\rho(h_1)d\rho(h_2)\\
=& \frac{1}{\rho(H_t)^2}\int_{H_t\times H_t} \left\langle \pi_Y(h_2^{-1}h_1)F, F\right\rangle \, d\rho(h_1)d\rho(h_2)\\
\le& \frac{1}{\rho(H_t)^2}\int_{H_t\times H_t} P(h_2^{-1}h_1)^{-\kappa}\|F\|^2_{L_l^2(B)} \, d\rho(h_1)d\rho(h_2).
\end{aligned}$$

We observe that since the polynomial P is proper on $\mathrm{Mat}_d(\mathbb{R})$, we have

$$\inf\{P(h) :\, h \in H\} > 0,$$

and there exists $a > 0$ such that

$$\rho(H_s) \ll e^{as}\,.$$

This simple estimate follows for an algebraic subgroup H from the much sharper asymptotic result stated in Theorem 8.1(i), but holds true for any closed subgroup H whatsoever.

Therefore, we deduce that

$$\begin{aligned}
&\int_{H_t\times H_t} P(h_2^{-1}h_1)^{-\kappa}\, d\rho(h_1)d\rho(h_2)\\
=&\int_{(h_1,h_2)\in H_t\times H_t : P(h_2^{-1}h_1)<e^s} P(h_2^{-1}h_1)^{-\kappa}\, d\rho(h_1)d\rho(h_2)\\
&+\int_{(h_1,h_2)\in H_t\times H_t : P(h_2^{-1}h_1)\ge e^s} P(h_2^{-1}h_1)^{-\kappa}\, d\rho(h_1)d\rho(h_2)\\
\le&\int_{(h,h_2)\in H\times H_t : P(h)<e^s} P(h)^{-\kappa}\, d\rho(h)d\rho(h_2) + e^{-\kappa s}\rho(H_t)^2\\
\ll&\, \rho(\{h\in H :\, P(h)<e^s\})\rho(H_t) + e^{-\kappa s}\rho(H_t)^2\\
\ll&\, e^{as}\rho(H_t) + e^{-\kappa s}\rho(H_t)^2.
\end{aligned}$$

Now taking s such that $e^s = \rho(H_t)^{1/(a+\kappa)}$, we conclude that for all $t \ge t_0$,

$$\|\pi_Y(\beta_t)F\|^2_{L^2(B)} \ll \rho(H_t)^{-\kappa/(a+\kappa)}\|F\|^2_{L_l^2(B)}, \quad F \in L_l^2(B).$$

This proves the theorem for $p = 2$, for the family $\pi_Y(\beta_t)$. The case of the family $\pi_Y(\beta_t^{g_1,g_2})$, as well as the uniformity as g_1, g_2 vary in compact sets in G is similar.

In order to complete the proof in general, we observe that the linear operator

$$\mathcal{A}_t(F) = \pi_Y(\beta_t)F - \int_Y F\, d\mu$$

satisfies the estimates

$$\begin{aligned}
&\|\mathcal{A}_t\|_{L_l^1(B)\to L^1(Y)} \ll 1,\\
&\|\mathcal{A}_t\|_{L_l^2(B)\to L^2(Y)} \ll \rho(H_t)^{-\kappa/(a+\kappa)},\\
&\|\mathcal{A}_t\|_{L_l^\infty(B)\to L^\infty(Y)} \ll 1.
\end{aligned}$$

Hence, the general case follows from Theorem 9.4 below. $\square$

Theorem 9.4 ([BS])**.** *Let $1 \le p_1 \le p_2 \le \infty$, $l \in \mathbb{N}$, B a compact domain in Y, and let*

$$\mathcal{A} : L_l^{p_1}(B) + L_l^{p_2}(B) \to L^{p_1}(Y) + L^{p_2}(Y)$$

be a linear operator such that

$$\mathcal{A}(L_l^{p_1}(B)) \subset L^{p_1}(Y), \quad \|\mathcal{A}\|_{L_l^{p_1}(B)\to L^{p_1}(Y)} \le M_1,$$
$$\mathcal{A}(L_l^{p_2}(B)) \subset L^{p_2}(Y), \quad \|\mathcal{A}\|_{L_l^{p_2}(B)\to L^{p_2}(Y)} \le M_2.$$

Then for every p such that $\frac{1}{p} = \frac{1-\theta}{p_1} + \frac{\theta}{p_2}$ with $\theta \in (0,1)$,

$$\mathcal{A}(L_l^{p}(B)) \subset L^{p}(Y), \quad \|\mathcal{A}\|_{L_l^{p}(B)\to L^{p}(Y)} \ll M_1^{1-\theta} M_2^{\theta}.$$

Proof. This theorem is a consequence for the results on interpolation of linear operators that can be found in [BS, Ch. 5]. Indeed, using a partition of unity, one can reduce the proof to the case when the support of $\phi \in L_l^p(B)$ is contained in a single coordinate chart. Then by the DeVore–Scherer Theorem (see [BS, Cor. 5.13]), the interpolation space $(L_l^1(\mathbb{R}^d), L_l^\infty(\mathbb{R}^d))_{1-1/p,p}$ is precisely $L_l^p(\mathbb{R}^d)$, and by the reiteration theorem [BS, Th. 5.12],

$$L_l^p(\mathbb{R}^d) = (L_l^{p_1}(\mathbb{R}^d), L_l^{p_2}(\mathbb{R}^d))_{\theta,p}$$

where θ is given by $\frac{1}{p} = \frac{1-\theta}{p_1} + \frac{\theta}{p_2}$. Similarly,

$$L^p(\mathbb{R}^d) = (L^{p_1}(\mathbb{R}^d), L^{p_2}(\mathbb{R}^d))_{\theta,p}$$

(see [BS, Th. 1.9]). Therefore, Theorem 9.4 is a consequence of [BS, Cor. 1.12]. □

9.3. **Quantitative pointwise ergodic theorem in Sobolev spaces.** Let us now note the following general fact based on the representation theory of a semisimple Lie group G, which applies to any of its closed subgroups, not just the algebraic ones.

Theorem 9.5. *Assume that*

- *G is a connected semisimple Lie group with finite centre acting smoothly on a manifold Y preserving a probability measure μ,*
- *the representation of every simple factor of G on $L_0^2(Y)$ is isolated from the trivial representation.*
- *H is an arbitrary closed subgroup of G, and the restricted sets H_t have exponential volume growth, namely $\rho(H_t) \gg e^{at}$ for some $a > 0$.*

Then there exist $l \in \mathbb{N}$ and $t_0 \ge 0$ such that

(i) Strong maximal inequality. *If the sets H_t satisfy the rough monotonicity property, that is, $\rho(H_{\lfloor t\rfloor+1}) \ll \rho(H_t)$, then for every $1 < p \le \infty$, a compact domain B of Y, and $F \in L_l^p(B)^+$,*

$$\left\| \sup_{t\ge t_0} \pi_Y(\beta_t) F \right\|_{L^p(Y)} \ll_{p,B} \|F\|_{L_l^p(B)}.$$

(ii) Quantitative pointwise theorem. *If the sets H_t satisfy in addition the Holder regularity property as stated in equation (8.3), then for every $1 < p < \infty$, a*

compact domain B of Y, a bounded function $F \in L^p_l(B)$, and almost every $y \in Y$,

$$\left|\pi_Y(\beta_t)F(y) - \int_Y F\,d\mu\right| \le C_p(y,F)e^{-\delta_p t}, \quad t \ge t_0,$$

where $\delta_p > 0$ and the estimator $C_p(y,F)$ satisfies

$$\|C_p(\cdot,F)\|_{L^p(Y)} \ll_{p,B} \|F\|_{L^p_l(B)} + \|F\|_\infty.$$

When H is a connected almost algebraic subgroup, the sets H_t are indeed roughly monotone and Hölder-regular, and the foregoing assertions apply to each family $\pi_Y(\beta_t^{g_1,g_2})$, uniformly as g_1, g_2 vary in compact sets in G.

Remark 9.6. We note that the maximal inequality just stated can be improved to an exponential-maximal inequality. The proof uses the analytic interpolation theorem, the estimates established below, and the interpolation results stated in Section §9.2, in a manner analogous to [MNS]. Such an argument renders superfluous the assumption in Theorem 9.5(ii) that the function is bounded.

Proof of Theorem 9.5(i). This estimate clearly holds for $p = \infty$, so let us assume that $p < \infty$. Since H_t have exponential growth, it follows from Theorem 9.3 that there exists $\delta_p > 0$ such that for every $F \in L^p_l(B)$,

$$\left\|\pi_Y(\beta_t)F - \int_Y F\,d\mu\right\|_{L^p(Y)} \ll e^{-\delta_p t}\|F\|_{L^p_l(B)}, \quad t \ge t_0.$$

Hence, for the function $C(y,F)$ defined by

$$C(y,F) := \sum_{n\in\mathbb{N}:n\ge t_0} \left|\pi_Y(\beta_n)F(y) - \int_Y F\,d\mu\right|,$$

we deduce from the triangle inequality that

$$\|C(\cdot,F)\|_{L^p(Y)} \ll \|F\|_{L^p_l(B)}.$$

Therefore,

$$\left\|\sup_{n\in\mathbb{N}:\,n\ge t_0}\left|\pi_Y(\beta_n)F - \int_Y F\,d\mu\right|\right\|_{L^p(Y)} \ll \|F\|_{L^p_l(B)}.$$

Then

$$\begin{aligned}\left\|\sup_{n\in\mathbb{N}:\,n\ge t_0}|\pi_Y(\beta_n)F|\right\|_{L^p(Y)} &\le \left\|\sup_{n\in\mathbb{N}:\,n\ge t_0}\left|\pi_Y(\beta_n)F - \int_Y F\,d\mu\right|\right\|_{L^p(Y)} + \left|\int_Y F\,d\mu\right| \\ &\ll \|F\|_{L^p_l(B)}.\end{aligned}$$

Finally, to complete the proof, we assume that $F \ge 0$. Then it follows from Theorem 8.1(i) that when H is a connected almost algebraic group $\rho(H_{\lfloor t\rfloor+1}) \ll \rho(H_t)$ for all $t \ge t_0$. Whenever this rough monotonicity property holds, we have

$$\begin{aligned}\pi_Y(\beta_t)F(y) &= \frac{1}{\rho(H_t)}\int_{H_t} F(h^{-1}y)\,d\rho(y) \le \frac{1}{\rho(H_t)}\int_{H_{\lfloor t\rfloor+1}} F(h^{-1}y)\,d\rho(y) \\ &\ll \pi_Y(\beta_{\lfloor t\rfloor+1})F(y).\end{aligned}$$

In conjunction with the previous estimate, this completes the proof. □

Proof of Theorem 9.5(ii). As already noted, it follows from Theorem 9.3 that for some $\delta > 0$,

$$\left\|\pi_Y(\beta_t)F - \int_Y F\,d\mu\right\|_{L^p(Y)} \ll e^{-\delta t}\|F\|_{L^p_l(B)}, \quad t \geq t_0. \tag{9.4}$$

We take an increasing sequence $\{t_i\}_{i\geq 0}$ that contains all positive integers greater than t_0 and has spacing $\lfloor e^{p\delta n/4}\rfloor^{-1}$ on the intervals $[n, n+1]$, $n \in \mathbb{N}$. Then

$$t_{i+1} - t_i \leq e^{-p\delta\lfloor t_i\rfloor/4}$$

for all $i \geq 0$. It follows from (9.4) that

$$\int_Y \left(\sum_{i\geq 0} e^{p\delta t_i/2}\left|\pi_Y(\beta_{t_i})F(y) - \int_Y F\,d\mu\right|^p\right) d\mu(y)$$
$$\ll \sum_{i\geq 0} e^{-p\delta t_i/2}\|F\|^p_{L^p_l(B)} \leq \sum_{n\geq\lfloor t_0\rfloor} e^{-p\delta n/2}\lfloor e^{p\delta n/4}\rfloor\|F\|^p_{L^p_l(B)} \ll \|F\|^p_{L^p_l(B)}.$$

Hence, if we set

$$C(y,F) := \left(\sum_{i\geq 0} e^{p\delta t_i/2}\left|\pi_Y(\beta_{t_i})F(y) - \int_Y F_{t_i}\,d\mu\right|^p\right)^{1/p},$$

then

$$\left|\pi_Y(\beta_{t_i})F(y) - \int_Y F\,d\mu\right| \leq C(y,F)e^{-p\delta t_i/2}$$

for all $i \geq 0$, and

$$\|C(\cdot,F)\|_{L^p(Y)} \ll \|F\|_{L^p_l(B)}.$$

For every $t \geq t_0$, there exists $t_i < t$ such that

$$t - t_i \ll e^{-p\delta\lfloor t_i\rfloor/4} \ll e^{-p\delta t/4}.$$

Then

$$\left|\pi_Y(\beta_t)F(y) - \int_Y F\,d\mu\right| \leq |\pi_Y(\beta_t)F(y) - \pi_Y(\beta_{t_i})F(y)|$$
$$+ \left|\pi_Y(\beta_{t_i})F(y) - \int_Y F\,d\mu\right|,$$

and the following computation completes the proof

$$
\begin{aligned}
&|\pi_Y(\beta_t)F(y) - \pi_Y(\beta_{t_i})F(y)| \\
=&\left|\frac{1}{\rho(H_t)}\int_{H_t} F(h^{-1}y)\,d\rho(y) - \frac{1}{\rho(H_{t_i})}\int_{H_{t_i}} F(h^{-1}y)\,d\rho(y)\right| \\
=&\left|\frac{1}{\rho(H_t)}\int_{H_t} F(h^{-1}y)\,d\rho(y) - \frac{1}{\rho(H_t)}\int_{H_{t_i}} F(h^{-1}y)\,d\rho(y)\right| \\
&+\left|\frac{1}{\rho(H_t)}\int_{H_{t_i}} F(h^{-1}y)\,d\rho(y) - \frac{1}{\rho(H_{t_i})}\int_{H_{t_i}} F(h^{-1}y)\,d\rho(y)\right| \\
\leq& 2\frac{\rho(H_t - H_{t_i})}{\rho(H_t)}\|F\|_\infty \ll (e^{-p\delta t/4})^\theta \|F\|_\infty,
\end{aligned}
$$

where the last estimate follows from our assumption that H_t is Hölder-regular. For a connected almost algebraic group H, Theorem 8.1(ii) shows that this property does indeed hold.

Finally, the proof for each family $\pi_Y(\beta_t^{g_1,g_2})$ supported on $H_t[g_1, g_2]$ is similar, and the uniformity as g_1, g_2 vary in compact sets in G follows from the uniform norm estimate established in Theorem 9.3. $\square$

For future reference below we also recall the following result about ergodic theory of semisimple groups which is a variation on [GN1, Th. 4.2]:

Theorem 9.7. *Assume that $H \subset G \subset \mathrm{SL}_d(\mathbb{R})$, G is an algebraic subgroup and P a homogeneous polynomial. Assume also that*

- *the group H is connected and semisimple, and it acts on the probability space (Y, μ) preserving the measure,*
- *the representation of every simple factor of H on $L_0^2(Y)$ is isolated from the trivial representation.*

Then, for the operators $\pi_Y(\beta_t)$ supported on the restricted sets H_t, we have for $t \geq t_0$,

(i) Strong exponential maximal inequality. *For every $1 < p < q \leq \infty$ and $F \in L^p(Y)$, with $\delta'_{p,q} > 0$*

$$\left\|\sup_{t\geq t_0} e^{\delta'_{p,q}t}|\pi_Y^0(\beta_t)F|\right\|_{L^p(Y)} \ll_p \|F\|_{L^q(Y)}\,.$$

(ii) Quantitative mean ergodic theorem. *For every $1 \leq p \leq q \leq \infty$ with $(p,q) \neq (1,1)$ and $(p,q) \neq (\infty,\infty)$, and $F \in L^q(Y)$, with $\delta_{p,q} > 0$*

$$\left\|\pi_Y(\beta_t)F - \int_Y F\,d\mu\right\|_{L^p(Y)} \ll_{p,q} e^{-\delta_{p,q}t}\|F\|_{L^q(Y)}\,.$$

Furthermore, the same conclusion holds without change for the operators $\pi_Y(\beta_t^{g_1,g_2})$ supported on $H_t[g_1, g_2]$, as g_1, g_2 vary over compact sets in G.

Proof. For any given family $\pi_Y(\beta_t^{g_1,g_2})$ the result is a direct consequence of [GN1, Th. 4.2]. The main ingredient in the proof is the strong spectral gap, which implies

the representation π_Y^0 on $L_0^2(Y)$ restricted to H is strongly L^v, for some $v < \infty$. Using the transfer principle, the Kunze-Stein phenomenon, and coarse admissibility of H_t, there is a uniform norm bound of the operators, as g_1, g_2 vary in a compact set, namely $\|\pi_Y^0(\beta_t^{g_1,g_2})\| \ll e^{-\kappa_v t}$, with $\kappa_v > 0$. This implies that the proof of [GN1, Th. 4.2] applies uniformly as g_1, g_2 vary in a compact set. □

9.4. **Spectral and volume estimates on non-amenable algebraic subgroups.** Let H be a closed almost connected subgroup of $\mathrm{SL}_d(\mathbb{R})$. We fix a non-negative proper homogeneous polynomial P on $\mathrm{Mat}_d(\mathbb{R})$ and set

$$H_t = \{h \in H : \log P(h) \le t\}.$$

We recall that β_t denote the Haar-uniform probability measure on H_t.

We denote by R the amenable radical of H, that is, the maximal closed connected normal amenable subgroup of H. We denote by ρ_R and ρ_H the corresponding Haar measures.

Definition 9.8. We say that H is *non-amenably embedded* (w.r.t. the gauge function P) if

$$\limsup_{t\to\infty} \frac{\log \rho_R(R \cap H_t)}{\log \rho_H(H_t)} < 1. \tag{9.5}$$

We note that if H is non-amenably embeded w.r.t. one homogeneous polynomial as above, then it is non-amenably embedded w.r.t. all of them, so that this notion is independent of the homogeneous polynomial chosen to verify it.

Of course, if H is non-amenably embedded (w.r.t. any gauge function P as above), then H is a non-amenable group. Let us recall the following definition [GN1, Ch.5]

Definition 9.9. *Groups with an Iwasawa decomposition.*

(i) An lcsc group H has an Iwasawa decomposition if it has two closed amenable subgroups K and Q, with K compact and $H = KQ$.
(ii) The *Harish-Chandra Ξ-function* associated with the Iwasawa decomposition $H = KQ$ of the unimodular group H is given by

$$\Xi_H(h) = \int_K \delta^{-1/2}(hk)dk$$

where δ is the left modular function of Q, extended to a left-K-invariant function on $H = KQ$. (Thus, if m_Q is left Haar measure on Q, $\delta(q)m_Q$ is right invariant, and $dm_H = dm_K \delta(q) dm_Q$.)

We begin by stating the following basic spectral estimates for Iwasawa groups, which follows from [GN1, Ch. 5, Prop. 5.9].

Theorem 9.10. *Let H be a unimodular lcsc group with an Iwasawa decomposition, and π a strongly continuous unitary representation of H. Assume that the sets H_t are coarsely admissible (i.e, satisfy condition (CA2) from §3.1). Then there exists $c > 0$ such that for every $t \ge t_0$, the following estimates hold.*

(i) *If π is weakly contained in the regular representation, and in particular if π is the regular representation* reg_H *itself, then*

$$\|\pi(\beta_t)\| \ll \frac{1}{\mathrm{vol}(H_{t+c})} \int_{H_{t+c}} \Xi_H(h)\, d\rho_H(h)\,.$$

(ii) *If $\pi^{\otimes 2N}$ is weakly contained in the regular representation of H, then*

$$\|\pi(\beta_t)\| \ll \|\mathrm{reg}_H(\beta_{t+c})\|^{\frac{1}{2N}}\ .$$

We now apply the previous general estimates to the case of subgroups of $\mathrm{SL}_d(\mathbb{R})$.

Proposition 9.11. *Let $H \subset \mathrm{SL}_d(\mathbb{R})$ be closed, unimodular and almost connected subgroup, and suppose that the group H is nonamenably embedded. Assume that the sets H_t are coarsely admissible (i.e, satisfy condition (CA2) from §3.1) and satisfy $\rho_H(H_t) \sim c_b\, e^{at} t^b$ as $t \to \infty$, with $c_b, a > 0$. Then*

(i) *For every $p > 0$ and $t \geq t_0$,*

$$\int_{H_t} \Xi_H(h)^p\, d\rho_H(h) \ll_p \rho_H(H_t)^{1-\delta_p},$$

where $\delta_p > 0$.

(ii) *The convolution norm of* $\mathrm{reg}_H(\beta_t)$ *as operators on $L^2(H)$ satisfies the decay estimate*

$$\|\mathrm{reg}_H(\beta_t)\|_{L^2(H)\to L^2(H)} \ll \mathrm{vol}(H_t)^{-\kappa}\,,$$

with $\kappa > 0$.

Furthermore, (ii) applies to each family $\pi_Y(\beta_t^{g_1,g_2})$ supported on $H_t[g_1, g_2]$, uniformly as g_1, g_2 vary in compact sets in G. In particular, the assertions above hold when H is an almost algebraic group which is non-amenably embedded in a semisimple group $G \subset \mathrm{SL}_d(\mathbb{R})$.

Proof. Since H has finitely many connected components, without loss of generality, we may assume that H is connected. Let $H = RL$ be the decomposition of H where L is a connected semisimple subgroup without compact factors, and R is the amenable radical. The existence of such a decomposition is an immediate consequence of the Levi decomposition. A left Haar measure ρ_H on H is given by product of a left Haar measure ρ_R on R and a Haar measure ρ_L on L. We can further identify an Iwasawa decomposition of H in the form $H = KQ$, where $K \subset L$ is a maximal compact subgroup, and $Q = Q_L R$, where $Q_L \subset L$ is a minimal parabolic subgroup of L.

When H is non-amenably embedded in G, by (9.5) there exists $\kappa > 0$ such that

$$\rho_R(R \cap H_t) \leq \rho_H(H_t)^{1-\kappa} \tag{9.6}$$

for all sufficiently large t.

By our assumption, there exist $c_b > 0$, $a \in \mathbb{Q}_{>0}$ and $b \in \mathbb{Z}_{\geq 0}$ such that

$$\rho_H(H_t) \sim c_b\, e^{at} t^b \quad \text{as } t \to \infty. \tag{9.7}$$

For a non-compact almost algebraic subgroup H this follows from Theorem 8.1.

Let $L(s) := \{l \in L;\, d(eK, lK) \le s\}$ denote the ball of radius s with respect to the Cartan-Killing metric on the symmetric space L/K of L. We use the estimate

$$\rho_L(L(s)) \ll e^{a's}, \quad s \ge 0, \tag{9.8}$$

with $a' > 0$.

Since the coordinates of the matrices in the set $L(s) = L(s)^{-1}$ are bounded by $e^{a''s}$ with some $a'' > 0$, up to a multiplicative constant, arguing as in Lemma 8.3, we conclude that there exists $c > 0$ such that for every $s > 0$,

$$H_t \cdot L(s)^{-1} \subset H_{t+cs}.$$

Let $\delta > 0$. Using (9.6), (9.7) and (9.8), we deduce that

$$\begin{aligned}\rho_H(RL(\delta t) \cap H_t) &\le \rho_H((R \cap H_{t+c\delta t})L(\delta t)) = \rho_R(R \cap H_{t+c\delta t})\rho_L(L(\delta t)) \\ &\le \rho_H(H_{t+c\delta t})^{1-\kappa}\rho_L(L(\delta t)) \ll e^{\theta t},\end{aligned}$$

where $\theta = (1+c\delta)(1-\kappa)a+\delta a'$ and t is sufficiently large. Hence, taking δ sufficiently small, we obtain $\theta < (1-\kappa/2)a$ and hence

$$\rho_H(RL(\delta t) \cap H_t) \le \rho_H(H_t)^{1-\kappa/2} \tag{9.9}$$

for all sufficiently large t.

To conclude the proof of Proposition 9.11, we recall the well-known estimate on the Harish-Chandra function Ξ_L on the semisimple Lie group L (see e.g. [GV, §4.6]): for every a in the positive Weyl chamber,

$$\Xi_L(a) \ll e^{-\rho_L(\log a)}(1 + \|\log a\|)^d,$$

where ρ_L denotes the half-sum of positive roots. Given $l \in L$, let its Cartan decomposition be given by $l = kak'$, where $k, k' \in K$ and a in the positive Weyl chamber. Note that

$$d(K, lK) = d(K, aK) = \|\log a\|.$$

Since there exists $\eta' > 0$ such that

$$\rho_L(\log a) \ge \eta' \, \|\log a\|$$

for a in the positive Weyl chamber, it follows that there exists $\eta > 0$ such that for $l \in L$,

$$\Xi_L(l) = \Xi_L(a) \ll e^{-\rho_L(\log a)}(1 + \|\log a\|)^d \ll e^{-\eta\, d(K,aK)} = e^{-\eta\, d(eK,lK)}. \tag{9.10}$$

Furthermore, the Harish-Chandra Ξ_H-function of the group H satisfies

$$\Xi_H(rl) = \Xi_H(lr) = \Xi_L(l)$$

for all $l \in L$ and $r \in R$. Indeed, every element in the amenable radical R acts trivially on the homogeneous space H/Q, since R is a normal subgroup of H contained in Q, so that $rhQ = hr'Q = hQ$. Hence,

$$\Xi_H(h) = \int_{H/Q} \sqrt{r_m(h, yQ)}\, dm(yQ),$$

where $r_m(h, yQ)$ is the Radon-Nikodym derivative of the unique K-invariant probability measure m on H/Q, which is invariant under left and right translations by $r \in R$.

Using bounds (9.9) and (9.10), we obtain that for all sufficiently large t,

$$\int_{RL(\delta t)\cap H_t} \Xi_H(h)^p \, d\rho_H(h) \ll_p \rho_H(H_t)^{1-\kappa/2},$$

and

$$\int_{H_t - RL(\delta t)} \Xi_H(h)^p \, d\rho_H(h) \ll_p e^{-p\eta\delta t}\rho_H(H_t).$$

This implies (i), and (ii) follows from Theorem 9.10. Clearly, when (CA1) is satisfied, the validity of the norm estimate stated in (ii) implies its validity for the families $H_t[g_1, g_2]$, uniformly as g_1, g_2 vary over compact sets in G. □

9.5. **Quantitative ergodic theorems for non-amenable algebraic subgroups.** We can now state the following general ergodic theorem in Lebesgue L^p-spaces, which applies in particular to algebraic subgroups H of a semisimple Lie group G.

Theorem 9.12. *Assume that $H \subset \mathrm{SL}_d(\mathbb{R})$ is a unimodular non-amenably embedded closed subgroup, and that the restricted sets H_t satisfy the Hölder property as in Theorem 8.1(ii) and have volume asymptotic $c_b t^b e^{at}$ with $c_b, a > 0$. Assume that $H \subset G$, where G is semisimple, and G acts on a probability space Y preserving an ergodic probability measure, such that the representation of G in $L_0^2(Y)$ has a strong spectral gap. Then*

(1) *The family $\pi_Y(\beta_t)$ satisfies the (L^p, L^r)-exponentially fast mean ergodic theorem for $1 < r \le p < \infty$, namely there exists $\delta_{p,r} > 0$ such that for every $f \in L^p(Y)$,*

$$\left\| \pi_Y(\beta_t) f - \int_Y f \, d\mu \right\|_{L^r(Y)} \ll_{p,r} e^{-\delta_{p,r} t} \|f\|_{L^p(Y)}$$

for all $t \ge t_0$.

(2) *The family $\pi_Y(\beta_t)$ satisfies the (L^p, L^r)-exponential strong maximal inequality for $1 < r < p < \infty$, namely there exist $t_0 > 0$ and $\delta_{p,r} > 0$ such that for every $f \in L^p(Y)$,*

$$\left\| \sup_{t \ge t_0} e^{\delta_{p,r} t} \left| \pi_Y(\beta_t) f - \int_Y f \, d\mu \right| \right\|_{L^r(Y)} \ll_{p,r} \|f\|_{L^p(Y)}.$$

(3) *The family β_t satisfies the (L^p, L^r)-exponentially fast pointwise ergodic theorem for $1 < r < p < \infty$, namely there exists $\delta_{p,r} > 0$ such that for every $f \in L^p(Y)$ and $t \ge t_0$,*

$$\left| \pi_Y(\beta_t) f(y) - \int_Y f \, d\mu \right| \le B_{p,r}(y, f) e^{-\delta_{p,r} t} \quad \text{for } \mu\text{-a.-e. } y \in Y$$

with the estimator $B_{p,r}(y, f)$ satisfying the norm estimate

$$\|B_{p,r}(\cdot, f)\|_{L^r(Y)} \ll_{p,r} \|f\|_{L^p(Y)}.$$

Furthermore, the same results hold without change for the operators $\pi_Y(\beta_t^{g_1,g_2})$ supported on the sets $H_t[g_1, g_2]$, uniformly when g_1, g_2 vary over compact sets in G.

In particular, the conclusions hold when H is an almost algebraic non-amenably embedded subgroup.

Proof. Since the representation π_Y^0 of G in $L_0^2(Y)$ has a strong spectral gap, it follows that for some even k, $(\pi_Y^0)^{\otimes k}$ is isomorphic to a subrepresentation of $\infty \cdot \mathrm{reg}_G$, i.e. to a subrepresentation of a multiple of the regular representation of G (see the discussion in [GN1, Ch. 5] for more details). It therefore follows that the restriction of π_Y^0 to the closed subgroup H has the same property, namely that $(\pi_Y^0|_H)^{\otimes k} \subset \infty \cdot \mathrm{reg}_H$. By Theorem 9.10 and Proposition 9.11 it follows that the exponentially fast mean ergodic theorem holds in L^2, uniformly for $\pi_Y(\beta_t^{g_1,g_2})$ as g_1, g_2 vary over compact sets in G. Using interpolation, it also holds as stated in part (i) of Theorem 9.12. The fact that under the regularity conditions stated in the Theorem, together with the norm decay established in part (i), the assertions of part (ii) and part (iii) follows is proved in detail in [GN1, Ch. 5] in the proof of Theorem 5.7. $\square$

10. Completion of the proof of the main theorems

We write the algebraic homogeneous space X of G as a factor space $X \simeq H\backslash G$ where H is an almost algebraic subgroup of G. The main theorems stated in the introduction will be deduced from the ergodic theory for the action of H on $Y \simeq G/\Gamma$ developed in Section 9 combined with the ergodic-theoretic duality results developed in Sections 3–7, and with the volume regularity properties established in Section 8.

10.1. **Regularity properties of the sampling sets.** We apply the results of Sections 3–8 to the sets

$$G_t = \{g \in G : \log P(g) \le t\} \quad \text{and} \quad H_t[g_1, g_2] = \{h \in H : \log P(g_1^{-1} h g_2) \le t\},$$

where P is either a non-negative proper homogeneous polynomial or norm on $\mathrm{Mat}_d(\mathbb{R})$. Let us verify that these sets satisfy the regularity properties used in Sections 3–7. When P is a non-negative proper homogeneous polynomial, we use the results established in Section 8. Property (CA1) follows from Lemma 8.3(i), and (6.1) of property (HA1) (and, in particular, (A1)) follows from Lemma 8.3(i). To complete verification of (HA1) we observe that a function χ_ε satisfying (6.2) can be constructed by identifying neighbourhoods of the identity in H with neighbourhoods of the origin in the Euclidean space and taking $\chi_\varepsilon(x) = \varepsilon^{-\dim(H)}\chi(\varepsilon x)$ for a fixed $\chi \in C_c^l(\mathbb{R}^{\dim(H)})$. This shows that (6.2) holds with $\kappa = (l + \dim(H)(1 - 1/q)$. Since property (6.3) follows from the corresponding property of the Euclidean space, we conclude that (HA1) holds. Property (CA2) follows from Theorem 8.1(i). Properties (A2), (A2′), (HA2), (H2′) with $1 \le r < \infty$ are established in Proposition 8.8. Property (A3) is a consequence of Proposition 8.5. When P is a norm, properties (CA1), (A1), and (HA1) directly follow from norm properties. The argument of [GN1, Prop. 7.3] gives the estimate

$$\rho(H_{t+\varepsilon}[u, v]) - \rho(H_t[u, v]) \le c\, \varepsilon \rho(H_t)$$

for all $t \ge t_0$ and $\varepsilon \in (0, 1)$, where c is uniform over u, v in compact sets. This implies conditions (CA1), (A2), (A2′), (HA2), (HA2′). Condition (A3) follows from the asymptotic formula for $\rho(H_t)$ established in [GW, Mau]. Finally, Properties (S) and (HS) are standard in the theory of homogeneous spaces of Lie groups.

Therefore, we conclude that the results established in Sections 3–7 apply in our setting.

10.2. **The limiting density.** It follows from Theorem 8.1(i) that for some $c_b > 0$, $a \in \mathbb{Q}_{\ge 0}$, and $b \in \mathbb{Z}_{\ge 0}$,

$$\rho(H_t) = c_b\, e^{at} t^b + O(e^{at} t^{b-1}). \tag{10.1}$$

We recall that by Lemma 8.2, $a = 0$ if and only the Zariski closure of H is an almost direct product of a compact subgroup and an abelian diagonalisable subgroup as in Theorem 1.1. The quantity $V(t) := e^{at} t^b$ is the correct normalisation for our averages.

Recall that we defined in (2.7))

$$d\nu_x(y) = \left(\lim_{t\to\infty} \frac{\rho(H_t[\mathsf{s}(x), \mathsf{s}(y)])}{\rho(H_t)}\right) d\xi(y)$$

Here the limit exists and is positive and continuous by Proposition 8.5, and the measure ξ is defined by (2.4). One can verify that the measures ν_x are canonically defined, i.e., they are independent of a choice of the section s and the Haar measure ρ on H.

We denote $\tilde{\nu}_x$, $x \in X$, the family of measures on X defined similarly by using the alternative normalization

$$d\tilde{\nu}_x(y) = \left(\lim_{t\to\infty} \frac{\rho(H_t[\mathsf{s}(x), \mathsf{s}(y)])}{e^{at} t^b}\right) d\xi(y). \tag{10.2}$$

so that $\tilde{\nu}_x = c_b\, \nu_x$.

If the group H is of subexponential type, then it follows from Proposition 8.6 that

$$\tilde{\nu}_x = c_b\, \xi, \tag{10.3}$$

and since H is unimodular, this gives the unique (up to scalar) G-invariant measure on X.

10.3. **Proof of Theorem 1.1, Theorem 1.2 and Theorem 1.3.** Let us now compare the averages

$$\pi_X(\tilde{\lambda}_t)\phi(x) = \frac{1}{e^{at} t^b} \sum_{\gamma \in \Gamma_t} \phi(x\gamma)$$

with the averages

$$\pi_X(\lambda_t)\phi(x) = \frac{1}{\rho(H_t)} \sum_{\gamma \in \Gamma_t} \phi(x\gamma)$$

which formed the subject of the discussion in Sections 3–7. It follows from (10.1) that

$$|\pi_X(\tilde{\lambda}_t)\phi| \ll |\pi_X(\lambda_t)\phi|.$$

Therefore, Theorem 1.1(i) follows from Theorem 3.1(ii) combined with Theorem 9.1(ii), Theorem 1.2(i) follows from Theorem 3.1(ii) combined with Theorem 9.5(i), and Theorem 1.3(i) follows from Theorem 3.1(ii) combined with Theorem 9.7(i), provide that the group H is connected. In general, its connected component H^0 has finite index in H, and we can apply the previous argument to the finite cover $H^0\backslash G$ which implies Theorem 9.7(i) for $H\backslash G$.

Similarly, by (10.1),

$$|\pi_X(\tilde{\lambda}_t)\phi(x) - c_b\,\pi_X(\lambda_t)\phi(x)| \ll t^{-1}|\pi_X(\lambda_t)\phi(x)|,$$

and by Theorem 3.1(i),

$$\|\pi_X(\tilde{\lambda}_t)\phi - c_b\,\pi_X(\lambda_t)\phi\|_{L^p(D)} \ll t^{-1}\|\phi\|_{L^p(D)}. \tag{10.4}$$

Hence, Theorem 1.1(ii) follows from Theorem 4.1 combined with Theorem 9.1(iii), and Theorem 1.1(iv) follows from Theorem 5.1 combined with Theorem 9.1(iii).

Combining Theorem 4.1 with Theorem 9.3, we deduce that Theorem 1.2(ii) holds for $\phi \in L^p_l(D)^+$ with $p > 1$. Since it sufficient to prove convergence for a dense family of functions (see the proof of Theorem 4.1), this implies the claim of Theorem 1.2(ii). Theorem 1.2(iii) follows from Theorem 5.1 combined with Theorem 9.5(ii).

To prove Theorem 1.1(iii), we observe that by Theorem 6.1 and Theorem 9.3, for some $\delta > 0$,

$$\left\|\pi_X(\lambda_t)\phi(x) - \pi_X(\lambda_t^G)\phi\right\|_{L^p(D)} \ll t^{-\delta}\|\phi\|_{L^q_l(D)}, \tag{10.5}$$

where

$$\begin{aligned}\pi_X(\lambda_t^G)\phi(x) &= \frac{1}{\rho(H_t)}\int_{G_t}\phi(\mathsf{p}_X(\mathsf{s}(x)g))\,dm(g)\\ &= \frac{1}{\rho(H_t)}\int_{(y,h)\colon \mathsf{s}(x)^{-1}h\mathsf{s}(y)\in G_t}\phi(\mathsf{p}_X(h\mathsf{s}(y)))\,d\rho(h)d\xi(y)\\ &= \int_X \phi(y)\frac{\rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])}{\rho(H_t)}\,d\xi(y).\end{aligned}$$

It follows from Proposition 8.6 that

$$\left|\pi_X(\lambda_t^G)\phi(x) - \int_X \phi\,d\xi\right| \ll t^{-1}\|\phi\|_{L^1(D)} \tag{10.6}$$

uniformly as x varies in compact sets. Therefore, combining estimates (10.4), (10.5), and (10.6), we deduce Theorem 1.1(iii).

Theorem 1.2(iv) is deduced from Theorem 7.1 combined with Theorem 9.3. It follows from Theorem 7.1 that for some $\delta > 0$ and almost every $x \in D$,

$$\left|\sum_{\gamma\in\Gamma_t}\phi(x\gamma) - \int_{G_t}\phi(xg)\,dm(g)\right| \ll_{\phi,x} e^{-\delta t}\rho(H_t) \ll e^{(a-\delta)t}t^b.$$

Since by Proposition 8.10,

$$\int_{G_t}\phi(xg)\,dm(g) = e^{at}\left(\sum_{i=0}^{b} c_i(\phi,x)t^i\right) + O_{\phi,x}\left(e^{(a-\delta)t}\right),$$

this implies Theorem 1.2(iv).

Theorem 1.3(ii) is deduced similarly from Theorem 7.1 and Theorem 9.7(ii), when H is connected. In general, we reduce the argument to the action on the space $H^0\backslash G$ which is a finite cover of X. This implies Theorem 1.3(ii) in general. Theorem 1.3(iii) is derived from Theorem 7.1 combined with Theorem 9.7 and Theorem 9.12 in the same manner as Theorem 1.2(iv), using that the norm estimates of $\pi_Y(\beta_t)$ are available in Lebesgue spaces, rather than just Sobolev spaces, uniformly as g_1, g_2 vary in a compact set.

10.4. **Proof of Theorem 1.4.** First, let us note that to prove each of the three statement in Theorem 1.4, it suffices to prove them for non-negative functions in the function space under consideration. Thus we can assume that the function ϕ is non-negative, when convenient.

We begin by proving the mean ergodic theorem stated in Theorem 1.4(i). By Theorem 9.7(ii), Theorem 6.1 (applying the case $l = 0$), and Theorem 9.12, we conclude that for some $\delta > 0$,

$$\left\|\pi_X(\lambda_t)\phi(x) - \pi_X(\lambda_t^G)\phi(x)\right\|_{L^p(D)} \ll e^{-\delta t}\|\phi\|_{L^q(D)}. \tag{10.7}$$

Since

$$\pi_X(\lambda_t^G)\phi(x) = \int_X \phi(y)\frac{\rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])}{\rho(H_t)}\,d\xi(y),$$

and the density of ν_x with respect to ξ is given by

$$\lim_{t\to\infty}\frac{\rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])}{\rho(H_t)} = \Theta(\mathsf{s}(x),\mathsf{s}(y)),$$

it suffices to estimate

$$\left|\pi_X(\lambda_t^G)\phi(x) - \int_D \phi\,d\nu_x\right| = \left|\int_X \phi(y)\left(\frac{\rho(H_t[\mathsf{s}(x),\mathsf{s}(y)])}{\rho(H_t)} - \Theta(\mathsf{s}(x),\mathsf{s}(y))\right)d\xi(y)\right|. \tag{10.8}$$

Since we assume that the homogeneous polynomial in question is a norm, H is semisimple, and the volume growth of H_t is purely exponential, we can appeal to [Mau, Thm. 3], where the following regularity property is established for the volume of $H_t[g_1, g_2]$:

$$\rho(H_t[g_1,g_2]) = c(g_1,g_2)e^{at} + O(e^{(a-\delta)t})\,,$$

with $\delta > 0$ independent of g_1, g_2, and $c(g_1, g_2)$ and the implied constant uniform as g_1, g_2 vary in compact sets in G. This immediately implies an exponential decay estimate of the kernel that appears in equation (10.8) and the quantitative mean ergodic theorem follows.

The exponential-maximal inequality, namely, the estimate for the quantity

$$\sup_{t\ge t_0} e^{\delta_{p,w}t}\left|\pi_X(\lambda_t)\phi(x) - \int_X \phi\,d\nu_x\right|$$

stated in Theorem 1.4(ii) follow from the exponential decay estimate just established on the norms $\left\|\pi_X(\lambda_t)\phi(x) - \int_X \phi\,d\nu_x\right\|_{L^p(D)}$. This follows from the same argument as already used in the proof of Theorem 7.1.

The exponentially fast pointwise ergodic theorem stated in Theorem 1.4(iii) follows directly from Theorem 7.1 (applying the case $l = 0$), together with the pointwise estimate arising from equation (10.8) using the volume asymptotics just cited.

Finally, to prove the statement in Remark 1.5 that when H is semisimple, but the growth is not necessarily purely exponential (namely $b > 0$), the arguments cited above of [GW, Mau] yield

$$\rho(H_t[g_1,g_2]) = c(g_1,g_2)e^{at}t^b + O(e^{at}t^{b-\delta})\,,$$

uniformly as g_1, g_2 vary in compact sets.

This implies a rate of decay estimate for the kernel that appears in equation (10.8), and repeating the foregoing arguments using this estimate, we deduce that the the quantitative mean and pointwise ergodic theorems with speed $t^{-\eta_p}$.

Remark 10.1. We note that the argument establishing the volume asymptotics used above in [GW] and [Mau] use the triangle inequality for norms in an essential way, as well as the explicit formulas for the invariant measure on semisimple groups. On the other hand, the argument with resolution of singularities used in Theorem 8.1 applies to general groups. A very similar estimate holds, but its uniformity in x, y is not clear. Whenever such uniformity is established, equation (10.8) will provide quantitative results as in Theorem 1.4.

11. Examples and applications

We now turn to discuss some examples in detail. We have formulated most of the results in general, but remind the reader that in examples 11.1, 11.4, 11.5 and 11.7 below, if we choose the homogeneous polynomial P to be norm with purely exponential volume growth of balls, then the stronger results of Theorem 1.4 apply.

11.1. **Quadratic surfaces.** We discuss the action on the de-Sitter space mentioned in the introduction (see (1.6)). We observe that the group $G := \mathrm{SO}_{d,1}(\mathbb{R})^0$ acts transitively on X, and $X \simeq H\backslash G$ where

$$H = \mathrm{Stab}_G(e_1) = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{SO}_{1,d-1}(\mathbb{R})^0 \end{pmatrix}.$$

When $d = 2$, the group H is a one-dimensional $\mathbb{R}$-diagonalisable subgroup. Hence, we are in the setting of Theorem 1.1 in this case. When $d \geq 3$, H is a simple almost algebraic group, and the representation of H on $L^2_0(G/\Gamma)$ is isolated from the trivial representation. We will now see that we are in fact in the setting of Theorem 1.4, namely for our choice of norm the volume growth is purely exponential.

Let us compute the normalization factor $V(t)$ and the limit measures. Let

$$K_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathrm{SO}_{d-1}(\mathbb{R}) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad B^+ = \left\{ b_s = \begin{pmatrix} \mathrm{id} & 0 & 0 \\ 0 & \cosh s & \sinh s \\ 0 & \sinh s & \cosh s \end{pmatrix} \right\}_{s \geq 0}.$$

Then we have the Cartan decomposition

$$H = K_0 B^+ K_0,$$

and a Haar measure with respect to this decomposition is given by

$$d\rho(k_1, s, k_2) = dk_1 \, (\sinh s)^{d-2} ds \, dk_2, \quad (k_1, s, k_2) \in K_0 \times \mathbb{R}_{\geq 0} \times K_0,$$

where dk_1 and dk_2 denote the probability Haar measures on K_0. Since

$$\|k_1 b_s k_2\| = \|b_s\| = e^s + O(1),$$

and it follows that (up to the factor 2^{2-d})

$$\rho(H_t) \sim \begin{cases} t & \text{when } d = 2, \\ e^{(d-2)t} & \text{when } d \geq 3. \end{cases}$$

This gives the normalization factor $V(t)$.

To compute the limit measure, we use that H is a symmetric subgroup of G. We have the Cartan decomposition for G with respect to H given by

$$G = HAK,$$

where

$$K = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{SO}_d(\mathbb{R}) \end{pmatrix} \quad \text{and} \quad A = \left\{ a_r = \begin{pmatrix} \cosh r & 0 & \sinh r \\ 0 & \mathrm{id} & 0 \\ \sinh r & 0 & \cosh r \end{pmatrix} \right\}_{r \in \mathbb{R}}.$$

We note that the A-component of this decomposition is unique, and when $r \neq 0$, the K-component is unique modulo K_0. Therefore, $X \simeq H \backslash G$ can be identified (up to measure zero) with $A \times K_0 \backslash K$. More explicitly, the identification is given by the polar coordinates (1.5). We use this identification to give the section $\mathsf{s} : X \to G$. A Haar measure on G with respect to the Cartan decomposition is given by

$$dm(h, r, \omega) = d\rho(h)\,(\cosh r)^{d-1} dr\, d\omega, \quad (h, r, \omega) \in H \times \mathbb{R} \times K_0 \backslash K,$$

where $d\omega$ denotes the probability Haar measure on $K_0 \backslash K$. We normalise m, so that $m(G/\Gamma) = 1$. It follows that the measure ξ appearing in (2.4) is equal up to a constant to

$$d\xi(r, \omega) = (\cosh r)^{d-1} dr\, d\omega, \quad (r, \omega) \in \mathbb{R} \times K_0 \backslash K.$$

In fact, this is a Haar measure on X. Using that the norm is K-invariant, and A commutes with K_0, we deduce that for $x_1 = e_1 a_{r_1} \omega_1$, $h = k_1 b_s k_2$, $x_2 = e_1 a_{r_2} \omega_2$, we have

$$\begin{aligned} \|\mathsf{s}(x_1)^{-1} h \mathsf{s}(x_2)\| &= \|\omega_1^{-1} a_{r_1}^{-1} k_1 b_s k_2 a_{r_2} \omega_2\| = \|a_{r_1}^{-1} b_s a_{r_2}\| \\ &= c(r_1, r_2) e^s + O_{r_1, r_2}(1), \end{aligned}$$

where

$$c(r_1, r_2) = \left(1 + (\sinh r_1)^2\right)^{1/2} \left(1 + (\sinh r_2)^2\right)^{1/2}.$$

This implies that

$$\lim_{t \to \infty} \frac{\rho(H_t[\mathsf{s}(x_1), \mathsf{s}(x_2)])}{V(t)} = c(r_1, r_2)^{-(d-2)},$$

and the limit measure is given by

$$d\nu_v(r, \omega) = \left(1 + v_d^2\right)^{-(d-2)/2} \left(1 + (\sinh r)^2\right)^{-(d-2)/2} (\cosh r)^{d-1} dr\, d\omega.$$

Since for $d \geq 3$ the group H is simple and has a spectral gap in G/Γ, and we have chosen a norm such that H_t has purely exponential volume growth, Theorem 1.4 applies. We conclude that for every $\phi \in L^p(X)$, $p > 1$, with compact support and almost every $v \in X$, the following quantitative pointwise convergence theorem holds :

$$\frac{1}{e^{(d-2)t}} \sum_{\gamma \in \Gamma_t} \phi(v\gamma) = \frac{c_d(\Gamma)}{(1 + v_d^2)^{(d-2)/2}} \int_X \phi(r, \omega) \frac{(\cosh r)^{d-1} dr\, d\omega}{(1 + (\sinh r)^2)^{(d-2)/2}} + O_{p,v,\phi}(e^{-\delta_p t}) \tag{11.1}$$

for some $c_d(\Gamma) > 0$, and with a fixed $\delta_p > 0$, independent of v and ϕ .

11.2. **Projective spaces.** We return to the action on the projective space discussed in the Introduction (see (1.4)). We observe that $\mathbb{P}^{d-1}(\mathbb{R})$ is a homogeneous space of $G := \mathrm{SL}_d(\mathbb{R})$, and $\mathbb{P}^{d-1}(\mathbb{R}) \simeq H\backslash G$ where

$$H := \begin{pmatrix} \star & 0 \\ \star & \star \end{pmatrix} \subset G$$

is the maximal parabolic subgroup of G. Since $G = HK$ with $K := \mathrm{SO}_d(\mathbb{R})$, a Haar measure on G is given by

$$\int_G f\, dm(g) = \int_{(K\cap H)\backslash K} \int_H f(hk)\, d\rho(h) d\xi(k), \quad f \in L^1(G),$$

where ρ is the left Haar measure on H and ξ a Haar measure on $(K \cap H)\backslash K$. We normalise ξ to be the probability measure and normalise ρ so that $m(G/\Gamma) = 1$. Then under the identification $\mathbb{P}^{d-1}(\mathbb{R}) \simeq (K \cap H)\backslash K$, the measure ξ is the measure appearing in (2.4). It follows from K-invariance of $\|\cdot\|$ and [DRS, Appendix 1] that

$$\rho(H_t) = m(G_t) \sim c\, e^{(d^2-d)t} \quad \text{as } t \to \infty$$

with $c > 0$. Hence, the correct normalisation factor in (1.4) is $V(t) = e^{(d^2-d)t}$. By (10.2), the limit measure is given by

$$\nu_v(u) = \Theta(v,u) d\xi(u), \quad u, v \in (K \cap H)\backslash K,$$

where

$$\Theta(v,u) = \lim_{t\to\infty} \frac{\rho(H_t[v,u])}{e^{(d^2-d)t}} = \Theta(e,e),$$

by the K-invariance of the norm. This completes verification of (1.4). It is clear that the above argument applies to other compact homogeneous spaces of $\mathrm{SL}_d(\mathbb{R})$ such as the Grassmann varieties and the flag variety.

The limit formula (1.4) but without an error estimate was obtained in [G1].

11.3. **Spaces of frames.** Let Γ be a lattice in $\mathrm{SL}_d(\mathbb{R})$ and $\Gamma_t = \{\gamma \in \Gamma : \log\|\gamma\| \le t\}$ denote the norm balls with respect to the standard Euclidean norm $\|\gamma\| = \left(\sum_{i,j=1}^d \gamma_{ij}^2\right)^{1/2}$. We consider the action of Γ on the space $X = \prod_{i=1}^k \mathbb{R}^d$ with $k < d$. We demonstrate that one can analyse the asymptotic distribution of the averages $\sum_{\gamma\in\Gamma_t} \phi(x\gamma)$ on X with a help of Theorem 1.2. This question was studied in [G2] (and for the two dimensional case in [Le1, No, LP, MW]). Although the method of [G2] allows to compute the asymptotics of $\sum_{\gamma\in\Gamma_t} \phi(x\gamma)$, it is not capable to give a rate of convergence. Theorem 1.2(iv) implies that for any nonnegative continuous subanalytic function $\phi \in L_l^1(\mathbb{R}^d)$ with compact support, and for almost every $v \in X$, there exists $\delta > 0$ such that

$$\frac{1}{e^{(d-1)(d-k)t}} \sum_{\gamma\in\Gamma_t} \phi(v\gamma) = \frac{c_{d,k}(\Gamma)}{\mathrm{vol}(v)^{d-1}} \int_X \phi(w)\, \frac{dw}{\mathrm{vol}(w)} + O_{\phi,v}(e^{-\delta t}), \qquad (11.2)$$

where $c_{d,k}(\Gamma) > 0$, $\mathrm{vol}(v)$ denotes the Euclidean volume of the k-dimensional parallelepiped spanned by the tuple of vectors in v, and dw denotes the measure on X which is the product of the Lebesgue measures on $\mathbb{R}^d$.

To deduce (11.2) from Theorem 1.2 we observe that the subset of X consisting of linearly independent vectors is a single orbit of the group $G := \mathrm{SL}_d(\mathbb{R})$ which has full measure on X. Therefore, up to measure zero $X \simeq H\backslash G$ where

$$H := \begin{pmatrix} id & 0 \\ \star & \star \end{pmatrix} \subset G,$$

and we are in the setting of Theorem 1.2. It remains to compute the normalisation factor $V(t)$ and the limit measure, and that has already been done in [G2] (see [G2, Theorem 3]).

11.4. **Dense projections.** Let $H \subset \mathrm{SL}_n(\mathbb{R})$ and $L \subset \mathrm{SL}_m(\mathbb{R})$ be connected semisimple groups, $G = H \times L$, and let Γ be an lattice in G such that its image under the natural projection map $\pi : G \to L$ is dense. We investigate the distribution of $\pi(\Gamma)$ in L. We fix Euclidean norms on $\mathrm{Mat}_n(\mathbb{R})$ and $\mathrm{Mat}_m(\mathbb{R})$ and the set

$$\Gamma_t = \{\gamma = (h, \ell) : \|(h, \ell)\| := \sqrt{\|h\|^2 + \|\ell\|^2} < e^t\}.$$

Let us assume that the representation of every simple factor of H on $L^2_0(G/\Gamma)$ is isolated from the trivial representation. This is known to be the case when G has no compact factors and also when Γ is a congruence subgroup (see [KS]). In this case, Theorem 1.3 implies pointwise almost sure convergence with respect to a Haar measure λ on L. Namely, there exist $a \in \mathbb{Q}_{>0}$, and $b \in \mathbb{Z}_{\geq 0}$ such that for every non-negative continuous subanalytic function ϕ on L with compact support and for almost every $x \in L$, we have the asymptotic expansion

$$\frac{1}{e^{at}t^b} \sum_{\gamma \in \Gamma_t} \phi(x\pi(\gamma)) = \int_L \phi \, d\lambda + \sum_{i=1}^{b} c_i(\phi, x) t^{-i} + O_{\phi,x}(e^{-\delta t}) \tag{11.3}$$

with $\delta > 0$.

To deduce formula (11.3) from Theorem 1.4, all we need to do is to identify the limit measure. We choose the section $\mathsf{s}(\ell) = (e, \ell)$. Then the measure ξ in (2.4) is equal to λ. Since for $h \in H$ and l_1, l_2 in a compact subset of H,

$$\|(e, \ell_1^{-1}) \cdot (h, e) \cdot (e, \ell_2)\| = \|(h, \ell_1^{-1}\ell_2)\| = \|h\| + O(1),$$

it follows that

$$\rho(H_t[\mathsf{s}(\ell_1), \mathsf{s}(\ell_2)]) \sim \rho(H_t) \quad \text{as } t \to \infty.$$

Hence, by (10.2), the limit measure is a Haar measure on L.

Quantitative equidistribution. Let us note that under our assumption here Theorem 1.4 holds as well, so that in particular, the quantitative mean ergodic theorem is valid. In [GN3] we apply this fact to $\tilde{\lambda}_t$ and derive that quantitative equidistribution holds in this case. Namely, for Hölder functions convergence holds for *every* $\ell \in L$ with a fixed rate, and with the implied constant uniform over ℓ in compact sets. When the volume growth is purely exponential, the rate of equidistribution is $e^{-\delta t}$, and otherwise the rate is $t^{-\eta}$. Previously, the problem of distribution of dense projections was investigated in [GW, Sec. 1.5.2], but the method of [GW] does not yield any error term.

11.5. **Values of quadratic form.** Let Q be a nondegenerate indefinite quadratic form in d variables with $d \geq 3$, signature (p, q). Given a tuple of vectors $v = (v_1, \ldots, v_d)$ in $\mathbb{R}^d$, we denote by $\bar{Q}(v)$ the corresponding Gram matrix:

$$\bar{Q}(v) := (Q(v_i, v_j))_{i,j=1,\ldots d} \in \mathrm{Mat}_d(\mathbb{R}).$$

We denote by $\mathcal{F}_d$ the set of unimodular frames, namely the set of d-tuples of vectors $v = (v_1, \ldots, v_d)$ in $\mathbb{R}^d$ satisfying $\det(v_1, \ldots, v_d) = 1$. Let $\mathcal{F}_d(\mathbb{Z})$ denote the subset of unimodular frames with integral coordinates.

Note that for any unimodular frame, the representation of the quadratic form Q as a matrix w.r.t. the frame has the same determinant, which we will denote by Δ. Thus for $v \in \mathcal{F}_d$, we have $\bar{Q}(v) \in \mathcal{Q}_{p,q}(\Delta)$ where $\mathcal{Q}_{p,q}(\Delta)$ denotes the set of symmetric matrices with signature (p, q) and determinant Δ. We also use the same notation for the corresponding set of quadratic forms. It is known that for almost all quadratic forms Q in the space of nondegenerate quadratic forms of given dimension, the set $\bar{Q}(\mathcal{F}_d(\mathbb{Z}))$ is dense in $\mathcal{Q}_{p,q}(\Delta)$. The distribution of $\bar{Q}(\mathcal{F}_d(\mathbb{Z}))$ in $\mathcal{Q}_{p,q}(\Delta)$ was investigated in [GW, Sec. 1.5.1], and here we show that the asymptotic formula from [GW] holds with an exponentially decaying error term for almost all quadratic forms $Q \in \mathcal{Q}_{p,q}(\Delta)$.

We observe that the group $G := \mathrm{SL}_d(\mathbb{R})$ acts transitively on $\mathcal{Q}_{p,q}(\Delta)$ by

$$x \mapsto {}^t g x g, \quad x \in \mathcal{Q}_{p,q}(\Delta),\ g \in G,$$

and the space $\mathcal{Q}_{p,q}(\Delta)$ can be identified with $H \backslash G$ where $H \simeq \mathrm{SO}_{p,q}(\mathbb{R})$. Moreover, with respect to this action,

$$\bar{Q}(\mathcal{F}_d(\mathbb{Z})) = x_Q \cdot \Gamma,$$

where x_Q denotes the matrix corresponding to Q, and $\Gamma = \mathrm{SL}_d(\mathbb{Z})$.

Using Theorem 7.1 we deduce that for $\phi \in L^s(D)^+$, $s > 1$, and some $\delta > 0$ (independent of ϕ and v) and for almost all Q

$$\sum_{v \in \mathcal{F}_d(\mathbb{Z}) \colon \sum_i \|v_i\|^2 < e^t} \phi(Q(v)) = \int_{v \in \mathcal{F}_d \colon \sum_i \|v_i\|^2 < e^t} \phi(Q(v))\, dm(v) + O_{s,\phi,Q}\left(e^{(p(q-1)-\delta)t}\right).$$

Here m denote the G-invariant measure on $\mathcal{F}_d \simeq G$, normalized so that $m(G/\Gamma) = 1$. The main term in the volume growth of the sets H_t has been shown in [GW, proof of [Cor. 1.3, pp. 104-106] to be given by $Be^{p(q-1)t}$ when $p < q$, where B is a suitable normalizing constant depending only of the group. Thus the error estimate is $\mathrm{vol}(H_t)e^{-\delta t}$.

When $p = q$, the volume growth of H_t has main term $Bte^{tp(p-1)}$ [GW], but nevertheless the error estimate in the preceding statement is again $\mathrm{vol}(H_t)e^{-\delta t} = te^{(p(p-1)-\delta)t}$, as follows from Theorem 7.1.

Note however that we have chosen a norm to define the sets H_t, the group H is simple (provided $(p, q) \neq (2, 2)$), and H has a spectral gap in $L^2(G/\Gamma)$. Therefore Theorem 1.4 applies in the present case. As already noted, the volume growth of H_t is purely exponential if and only if $p \neq q$, and in the latter case the normalized sampling operators λ_t converge exponentially fast in L^2-norm and pointwise almost everywhere to the integral of ϕ w.r.t. the limiting density, for every $\phi \in L^s(\mathcal{Q}_{p,q}(\Delta))$, $s > 1$, with compact support and almost every $Q \in \mathcal{Q}_{p,q}(\Delta)$,

11.6. **Affine actions of solvable groups.** We now turn to discuss the affine action (1.1) mentioned in the introduction. It is clear that ergodicity of this action is equivalent to ergodicity of the action of the matrix a on the torus $\mathbb{R}^d/\Delta$. Hence, this action is ergodic if and only if the matrix a has no roots of unity as eigenvalues.

To see that Theorem 1.1 applies to this case, let us first consider the case when all the eigenvalues of the matrix a are positive. Then a can embedded in a one-parameter algebraic subgroup H of $\mathrm{SL}_d(\mathbb{R})$. We consider the exponential solvable group $G := H \ltimes \mathbb{R}^d$ which is naturally an algebraic subgroup $\mathrm{SL}_{d+1}(\mathbb{R})$ and contains Γ as a lattice. Then $\mathbb{R}^d \simeq H\backslash G$, and the sets Γ_t defined in (1.2) are given by $\Gamma_t = \{\gamma \in \Gamma : \log\|\gamma\|' \le t\}$ with respect to a suitable chosen norm $\|\cdot\|'$ on $\mathrm{Mat}_{d+1}(\mathbb{R})$. Hence, we are in the framework of Theorem 1.1, which holds for the normalized sampling operators supported on the sets defined by general norms (see Remark 8.7).

We now compute the normalization factor $V(t)$ and the limit measures ν_v, $v \in \mathbb{R}^d$, following the general formulas from Section 10. Let ρ be the Haar measure on H for which $\rho(H/\langle a\rangle) = 1$. Then

$$\rho(H_t) \sim t \quad \text{as } t \to \infty.$$

Hence, the correct normalization factor is $V(t) = t$. The measure

$$dm(h, x) = d\rho(h)\frac{dx}{\mathrm{vol}(\mathbb{R}^d/\Delta)}, \quad (h, v) \in H \ltimes \mathbb{R}^d,$$

is the Haar measure on G such that $m(G/\Gamma) = 1$. For the section $\mathsf{s}(x) = (e, x)$ of the factor map $G \to \mathbb{R}^d \simeq H\backslash G$, the corresponding measure ξ, defined by (2.4), is $d\xi(x) = \frac{dx}{\mathrm{vol}(\mathbb{R}^d/\Delta)}$, and according to (10.3) it is equal to the limit measure appearing in (1.3).

Finally, in the case when the matrix a has negative real eigenvalues one can apply the previous argument to the index 2 index subgroup $\langle a^2\rangle \ltimes \Delta$ of Γ to verify the claim.

We remark that another interesting collection of examples for which Theorem 1.1 applies arises in the case of dense subgroups of nilpotent groups. For a different approach to equidistribution results for dense nilpotent groups we refer to [Br].

11.7. **Affine actions of lattices.** Consider the affine action of the group $\Gamma = \mathrm{SL}_d(\mathbb{Z}) \ltimes \mathbb{Z}^d$ on the Euclidean space $\mathbb{R}^d$. It is natural to consider Γ as a subgroup of $\mathrm{SL}_{d+1}(\mathbb{R})$. We fix a Euclidean norm $\mathrm{Mat}_{d+1}(\mathbb{R})$ and define the sets Γ_t with respect to this norm. As we shall verify, Theorem 1.4 applies to the normalized sampling operators $\sum_{\gamma\in\Gamma_t}\phi(x\gamma)$ on $\mathbb{R}^d$. Therefore, we deduce that for every $\phi \in L^p(\mathbb{R}^d)$ of compact support, for almost every $\in \mathbb{R}^d$, and for a fixed $\delta_p > 0$ independent of ϕ and x,

$$\frac{1}{e^{(d^2-d)t}}\sum_{\gamma\in\Gamma_t}\phi(v\gamma) = \frac{c_d}{(1+\|v\|^2)^{d/2}}\int_{\mathbb{R}^d}\phi(x)\,dx + O_{p,\phi,x}\left(e^{-\delta_p t}\right), \tag{11.4}$$

where $c_d = \pi^{d^2/2}\Gamma(d/2)^{-1}\Gamma((d^2-d+2)/2)^{-1}\zeta(2)^{-1}\cdots\zeta(d)^{-1}$.

To verify (11.4), we consider the group Γ as a lattice subgroup in the group $G := \mathrm{SL}_d(\mathbb{R}) \ltimes \mathbb{R}^d$, which is naturally an algebraic subgroup of $\mathrm{SL}_{d+1}(\mathbb{R})$ under the

embedding

$$(h,x) \mapsto \begin{pmatrix} h & 0 \\ x & 1 \end{pmatrix}, \quad (h,x) \in G.$$

Then $\mathbb{R}^d$ is a homogeneous space of G with respect to the action by the affine transformations, and $\mathbb{R}^d \simeq H\backslash G$ where $H = \mathrm{SL}(d,\mathbb{R})$. The action $\mathrm{SL}_d(\mathbb{Z})$ on the torus $\mathbb{R}^d/\mathbb{Z}^d$ has spectral gap. Since the action of H on G/Γ is isomorphic to the action induced from this action, it follows that it has spectral gap as well. Thus H is simple and acts with a spectral gap on $L^2(G/\Gamma)$, we have defined the sets H_t using a norm, and we will see below that the rate of growth of H_t is purely exponential. The assumption of Theorem 1.4 are therefore satisfied.

It remains to compute the formulas for the normalization factor $V(t)$ and the limit measure ν_v (following the recipe of Section 10). We fix a Haar measures ρ on H such that $\rho(H/\mathrm{SL}_d(\mathbb{Z})) = 1$. Then

$$dm(h,x) = d\rho(h)dx, \quad (h,x) \in G,$$

is the Haar measure on G such that $m(G/\Gamma) = 1$. Hence, if we take the section $\mathsf{s} : \mathbb{R}^d \to G$ to be $\mathsf{s}(x) = (e,x)$, the measure ξ, defined by (2.4), is the Lebesgue measure on $\mathbb{R}^d$. Since

$$\rho(H_t) \sim c_d\, e^{(d^2-d)t} \quad \text{as } t \to \infty. \tag{11.5}$$

(see [DRS, Appendix 1]), it follows that the normalization factor should be $V(t) = e^{(d^2-d)t}$. By (10.2), the limit measure is given by

$$\nu_v(x) = \Theta((e,v),(e,x))dx,$$

where

$$\Theta((e,v),(e,x)) = \lim_{t\to\infty} \frac{\rho(H_t[(e,v),(e,x)])}{e^{(d^2-d)t}}.$$

We have

$$\begin{aligned} \rho(H_t[(e,v),(e,x)]) &= \rho(\{h \in H :\, \log \|(e,v)^{-1}\cdot(h,0)\cdot(e,x)\| < t\}) \\ &= \rho(\{h \in H :\, \|(h,-vh+x)\| < e^t\}) \\ &= \rho(\{h \in H :\, (\|h\|^2 + \|vh - x\|^2)^{1/2} < e^t\}). \end{aligned}$$

By the triangle inequality,

$$\|(h,-vh)\| - \|x\| \le \|(h,-vh+x)\| \le \|(h,-vh)\| + \|x\|.$$

This implies that the above limit is independent of x. Moreover, since the norm is invariant under $k \in \mathrm{SO}_d(\mathbb{R})$, we obtain that

$$\begin{aligned} \rho(\{h \in H :\, (\|h\|^2 + \|vkh\|^2)^{1/2} < e^t\}) &= \rho(\{h \in H :\, (\|k^{-1}h\|^2 + \|vh\|^2)^{1/2} < e^t\}) \\ &= \rho(\{h \in H :\, (\|h\|^2 + \|vh\|^2)^{1/2} < e^t\}). \end{aligned}$$

Therefore,

$$\begin{aligned} &\rho(\{h \in H :\, (\|h\|^2 + \|vh\|^2)^{1/2} < e^t\}) \\ =&\rho\left(\left\{ h \in H :\, \left(\sum_{i=1}^{d-1} \|e_i h\|^2 + (1+\|v\|^2)\|e_d h\|^2\right)^{1/2} < e^t \right\}\right) \end{aligned}$$

where $\{e_i\}_{i=1}^d$ is the standard basis of $\mathbb{R}^d$. Let

$$h_v = \mathrm{diag}\left(1, \dots, 1, (1+\|v\|^2)^{1/2}\right) = (1+\|v\|^2)^{1/(2d)} h'_v \in \mathrm{GL}_d(\mathbb{R}).$$

Then since $h'_v \in H = \mathrm{SL}_d(\mathbb{R})$, we get

$$\begin{aligned}\rho(\{h \in H:\ (\|h\|^2 + \|vh\|^2)^{1/2} < e^t\}) &= \rho(\{h \in H:\ \|h_v h\| < e^t\})\\ &= \rho(\{h \in H:\ \|h\| < (1+\|v\|^2)^{-1/(2d)} e^t\})\\ &\sim c_d\,(1+\|v\|^2)^{-(d-1)/(2)}\, e^{(d^2-d)t}\end{aligned}$$

as $t \to \infty$, by (11.5). This explains the formula for the limit measure in (11.4).

We also note that using the method of [GW], which is based on the Ratner's theory of unipotent flows, one can prove that

$$\lim_{t\to\infty} \frac{1}{e^{(d^2-d)t}} \sum_{\gamma \in \Gamma_t} \phi(v\gamma) = \frac{c_d}{(1+\|v\|^2)^{d/2}} \int_{\mathbb{R}^d} \phi(x)\,dx$$

for every irrational $v \in \mathbb{R}^d$.

References

[Aa] J. Aaronson, An introduction to infinite ergodic theory. Mathematical Surveys and Monographs, 50. American Mathematical Society, Providence, RI, 1997.

[AAB] C. Anantharaman-Delaroche, J-P. Anker, M. Babillot, et al, Théorèmes ergodiques pour les Action de Groupes. Monographie de L'Enseignement Mathematique, 41, 2010.

[Ar] V. Arnol'd, Arnold's problems. Springer-Verlag, Berlin, 2004.

[AK] V. Arnol'd and A. Krylov, Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain. Dokl. Akad. Nauk SSSR 148 (1963), 9–12.

[AGH] L. Auslander, L. Green and F. Hahn, Flows on homogeneous spaces. Annals of Math. Studies, vol. 53, 1963, Princeton University Press.

[BO] Y. Benoist and H. Oh, Effective equidistribution of S-integral points on symmetric varieties. arXiv:0706.1621.

[BS] C. Bennett and R. Sharpley, Interpolation of operators. Pure and Applied Mathematics, 129. Academic Press, Inc., Boston, MA, 1988.

[Br] E. Breuillard, Equidistribution of dense subgroups on nilpotent Lie groups. Ergodic Theory Dynam. Systems 30 (2010), 131–150.

[D1] S. G. Dani, Continuous equivariant images of lattice-actions on boundaries. Ann. of Math. (2) 119 (1984), 111–119.

[D2] S. G. Dani, Divergent trajectories of flows on homogeneous spaces and Diophantine approximation. J. Reine Angew. Math. 359 (1985), 55–89; erratum: J. Reine Angew. Math. 359 (1985), 214.

[DM] S. G. Dani and G. Margulis, Limit distributions of orbits of unipotent flows and values of quadratic forms. I. M. Gelfand Seminar, 91–137, Adv. Soviet Math., 16, Part 1, Amer. Math. Soc., Providence, RI, 1993.

[DR] S. G. Dani and S. Raghavan, Orbits of Euclidean frames under discrete linear groups. Israel Journal of Math 36 (1980), 300–320.

[DRS] W. Duke, Z. Rudnick, P. Sarnak, Density of integer points on affine homogeneous varieties. Duke Math. J. 71 (1993), 143–179.

[EMM] A. Eskin, G. Margulis and S. Mozes, Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture. Ann. of Math. (2) 147 (1998), 93–141.

[EM] A. Eskin and C. McMullen, Mixing, counting, and equidistribution in Lie groups. Duke Math. J. 71 (1993), 181–209.

[EMS] A. Eskin, S. Mozes and N. Shah, Unipotent flows and counting lattice points on homogeneous varieties. Ann. of Math. (2) 143 (1996), 253–299.

[F] H. Furstenberg, The unique ergodicity of the horocycle ow, in Recent Advances in Topological Dynamics (New Haven, Conn., 1972), Lecture Notes in Math. 318, Springer, Berlin, 1973, 95–115.

[GV] R. Gangolli and V. S. Varadarajan, Harmonic Analysis of Spherical Functions on Real Reductive Groups. Modern Surveys in Mathematics, 101, Springer Verlag, 1988.

[GGN] A. Ghosh, A. Gorodnik and A. Nevo, Diophantine approximation and automorphic spectrum, arXiv:1007.0593.

[G1] A. Gorodnik, Lattice action on the boundary of $\mathrm{SL}(n,\mathbb{R})$. Ergodic Theory Dynam. Systems 23 (2003), 1817–1837.

[G2] A. Gorodnik, Uniform distribution of orbits of lattices on spaces of frames. Duke Math. J. 122 (2004), 549–589.

[GM] A. Gorodnik and F. Maucourant, Proximality and equidistribution on the Furstenberg boundary. Geom. Dedicata 113 (2005), 197–213.

[GN1] A. Gorodnik and A. Nevo, The ergodic theory of lattice subgroups. Annals of Mathematics Studies 172, Princeton University Press, 2010.

[GN2] A. Gorodnik and A. Nevo, Counting lattice points. J. Reine Angew. Math. 663 (2012), 127–176.

[GN3] A. Gorodnik and A. Nevo, On Arnold's and Kazhdan's equidistribution problems, to appear in Ergodic Theory Dynam. Systems.

[GO] A. Gorodnik and H. Oh, Orbits of discrete subgroups on a symmetric space and the Furstenberg boundary. Duke Math. J. 139 (2007), 483–525.

[GW] A. Gorodnik and B. Weiss, Distribution of lattice orbits on homogeneous varieties. Geom. Funct. Anal. 17 (2007), 58–115.

[K] D. Kazhdan, Uniform distribution on a plane. Trudy Moskov. Mat. Obsh. 14 (1965), 299–305.

[KS] D. Kelmer and P. Sarnak, Strong spectral gaps for compact quotients of products of $\mathrm{PSL}(2,\mathbb{R})$. J. Eur. Math. Soc. 11 (2009), 283–313

[KM1] D. Kleinbock and G. Margulis, Bounded orbits of nonquasiunipotent flows on homogeneous spaces. Sinai's Moscow Seminar on Dynamical Systems, 141–172, Amer. Math. Soc. Transl. Ser. 2, 171, Amer. Math. Soc., Providence, RI, 1996.

[KM2] D. Kleinbock and G. Margulis, Flows on homogeneous spaces and Diophantine approximation on manifolds. Ann. of Math. (2) 148 (1998), 339–360.

[La] S. Lang, Report on diophantine approximations. Bull. Soc. Math. France 93 (1965), 177–192.

[Le1] F. Ledrappier, Distribution des orbites des rśeaux sur le plan réel. C. R. Acad. Sci. Paris Sér. I Math. 329 (1999), 61–64.

[Le2] F. Ledrappier, Ergodic properties of some linear actions, Journal of Mathematical Sciences 105, (2001), 1861–1875.

[LP] F. Ledrappier and M. Pollicott, Ergodic properties of linear actions of (2×2)-matrices. Duke Math. J. 116 (2003), 353–388.

[LP2] F. Ledrappier and M. Pollicott, Distribution results for lattices in $\mathrm{SL}(2,\mathbb{Q}_p)$. Bull. Braz. Math. Soc. (N.S.) 36 (2005), 143–176.

[Ma1] G. Margulis, On some aspects of the theory of Anosov systems. With a survey by Richard Sharp: Periodic orbits of hyperbolic flows. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2004.

[Ma2] G. Margulis, Oppenheim conjecture. Fields Medallists' lectures, 272–327, World Sci. Ser. 20th Century Math., 5, World Sci. Publ., River Edge, NJ, 1997.

[MNS] G. Margulis, A. Nevo and E. Stein, Analogs of Wiener's ergodic theorems for semi-simple Lie groups II. Duke J. Math. 103 (2000), 233–259.

[MS1] J. Marklof and A. Strömbergsson, The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems, Ann. Math. 172 (2010) 1949–2033.

[MS2] J. Marklof and A. Strömbergsson, The Boltzmann-Grad limit of the periodic Lorentz gas, Ann. Math. 174 (2011) 225–298.

[Mo] C. C. Moore, Ergodicity of flows on homogeneous spaces. Amer. J. Math. 88 (1966), 154–178.

[Mau] F. Maucourant, Homogeneous asymptotic limits of Haar measures of semisimple linear groups and their lattices. Duke Math. J. 136 (2007), 357–399.

[MW] F. Maucourant and B. Weiss, Lattice actions on the plane revisited, arXiv:1001.4924.

[Ne] A. Nevo, Pointwise ergodic theorems for actions of groups. Handbook of dynamical systems. Vol. 1B, 871–982, Elsevier B. V., Amsterdam, 2006.

[No] A. Nogueira, Orbit distribution on $\mathbb{R}^2$ under the natural action of $\mathrm{SL}(2,\mathbb{Z})$. Indag. Math. (N.S.) 13 (2002), 103–124.

[P] A. Parusinski, Subanalytic functions. Trans. Amer. Math. Soc. 344 (1994), 583–595.

[R] W. Rudin, Real and complex analysis. Third edition. McGraw-Hill Book Co., New York, 1987.

[Sh] N. Shah, Limit distributions of expanding translates of certain orbits on homogeneous spaces. Proc. Indian Acad. Sci. Math. Sci. 106 (1996), 105–125.

[Sp] R. Spatzier, On lattices acting on boundaries of semisimple groups. Ergodic Theory Dynamical Systems 1 (1981), 489–494.

[V1] W. A. Veech, Minimality of horospherical flows. Israel J. Math. 21 (1975), 233–239.

[V2] W. A. Veech, Unique ergodicity of horospherical flows. Amer. J. Math. 99 (1977), 827–859.

[Z1] R. Zimmer, Induced and amenable actions of Lie groups. Ann. Sci. Ec. Norm. Sup. 11 (1976), 407–428.

[Z2] R. Zimmer, Equivariant images of projective space under the action of $\mathrm{SL}(n,\mathbb{Z})$. Ergodic Theory Dynamical Systems 1 (1981), 519–522.

School of Mathematics, University of Bristol, Bristol, U.K.
E-mail address: `a.gorodnik@bristol.ac.uk`

Department of Mathematics, Technion, Israel
E-mail address: `anevo@tx.technion.ac.il`